Does the Consolidated Feed Matter?

Shihao Yu*

Current version: November 16, 2021

Abstract

In this paper I examine the role of the consolidated feed in the current US equities market by studying exogenous events affecting its speed and availability respectively. I find that while a faster consolidated feed has no impact on overall market liquidity, it leads to an increase in low-latency trading activity, possibly resulting from more competition between fast and slow traders as the latter becomes faster. Moreover, when the consolidated feed becomes corrupted or unavailable due to a technical glitch, market liquidity significantly worsens. My findings suggest that the consolidated feed remains a crucial component of today's market data infrastructure.

*School of Business and Economics, Vrije Universiteit Amsterdam. Email: s.yu@vu.nl. I thank Albert J. Menkveld, Bart Yueshen Zhou for their helpful suggestions and discussions. I gratefully acknowledge Netherlands Organisation for Scientific Research (NWO) for a Research Talent grant.

1 Introduction

Trading in the US equities market has become increasingly complicated. First of all, trading is highly fragmented and can now happen on any of the 16 public exchanges and dozens of broker-dealer internalizers and dark pools. Moreover, trading is extremely fast and happens at sub-millisecond frequencies.[1] In such a fragmented and fast trading environment, having access to low-latency market data is crucial for various market participants to implement their trading strategies: Market makers need fast data to constantly re-price their quotes and avoid being adverse-selected; Institutional investors need it to find the best available prices to trade and arbitragers need it to exploit short-lived arbitrage opportunities.

There are mainly two types of market data being used by traders in today's US equities market: consolidated feeds disseminated by so called security information processors (SIPs) and direct feeds disseminated by exchanges. While consolidated feeds are mandated by regulation, direct feeds are expensive proprietary products of the exchanges and normally have a lower latency and contain more information than consolidated feeds.[2] So some market participants argue that the discrepancy between the two types of feeds has created an unfair "two-tiered" market with the haves (direct feeds) and have not (consolidated feeds). Perhaps based on such reasoning, the SEC adopted a new rule named Reg NMS II in February 2021, aiming to further improve the consolidated feed and make it more comparable to direct feeds and thus level the playing field of market data. However, the rule has caused a fierce backlash from exchanges[3] which claim that further enhancements to the consolidated feed will not benefit the market.[4] Across the Atlantic, ESMA is scheduled to roll out a European-wide consolidated tape in the foreseeable future. Details over what should be included (only post-trades or including pre-trade quotes) and reporting latency is still very much in debate.

Objective. One argument by the exchanges that an enhanced consolidated feed will not benefit

[1]For example, Menkveld (2018) analyzes a sample of Nasdaq trades in October 2010 and finds that twenty percent of trades arrive in sub-millisecond clusters.

[2]See Section2.3.1 below for details about the speed and information difference between the consolidated feeds and direct feeds.

[3]SEC approved the rule in December, 2020 but was sued by Nasdaq, NYSE and Cboe. Thus the implementation of the rule is now blocked. See "Nasdaq, NYSE Sue SEC to Block Market Data Overhaul.", *The Wall Street Journal*, February 9, 2021.

[4]"Building the SIP Autobahn", *Nasdaq*, May 20, 2021.

the market is that it is primarily used by "display users" or human traders. Thus further speed improvement in the sub-second frequencies seems unlikely to benefit them. As for latency-sensitive traders, they will always opt for faster direct feeds and be hardly affected either. However, anecdotal evidence shows that the consolidated feed is more used by algorithmic traders as it becomes faster.[5] Moreover, there are scenarios where the consolidated feed can be useful to market participants. For example, regulatory filings show that about 45% dark pools do not use any direct feeds but consolidated feeds instead.[6] In addition, even for low-latency traders, they normally use data feeds from multiple sources for data integrity check (CFTC and SEC, 2010) and might withdraw from their trading when the consolidated feed becomes unreliable (Aldrich, 2017). A recent paper by Ernst, Sokobin, and Spatt (2021) shows that fast traders react to off-exchange trade reports from the consolidated feed even though they appear with a significant delay. So the key empirical question is: Does the consolidated feed really matter? If so, what role does it play in today's market? My objective in this paper is trying to shed some lights on these two questions.

Approach. To identify the role of the consolidated feed, I exploit events which exogenously affect its speed and availability respectively. The first event I use is a speed upgrade to the SIP operated by Nasdaq on October 24, 2016, which significantly reduced its median *processing* latency from about 350 microseconds to less than 20 microseconds. In the US equities market, there are two SIPs operated by Nasdaq and NYSE respectively, with the former being responsible for disseminating consolidated feeds for Nasdaq-listed stocks (Tape C securities) and the latter for NYSE-listed stocks (Tape A securities) and stocks listed on other regional exchanges and their successors (Tape B securities). Such a unique structure of two SIPs allows me to perform a difference-in-difference (DiD) analysis based on a matched sample of Nasdaq-listed and NYSE-listed stocks. If the speed of the consolidated feed matters for traders, I should see metrics of interest change in Nasdaq-listed stocks after the upgrade, *relative to* NYSE-listed stocks. To have an overall assessment of the potential impact, I compute a wide range of market liquidity measures and trading metrics from several databases including the NYSE's TAQ database, SEC's MIDAS database and LOBSTER database. In addition, I exploit the unique geographical locations of the exchanges and two-SIPs for further identification. Specifically, Nasdaq exchanges and Nasdaq-SIP are both located at Carteret,

[5] "Consolidated Market Data Feeds Gain Traction in Algo Trading and Fixed Income", *Finextra*, January 2019.
[6] See, for example, "Dispelling the Complementary Product Theory for Market Data", *Nasdaq*, August 8, 2020.

New Jersey while NYSE exchanges and NYSE-SIP are located about 35 miles away at Mahwah, New Jersey. For a trader co-located to Nasdaq exchanges trading Nasdaq-listed stocks, the speed advantage of direct feeds over consolidated feeds diminishes the most after the upgrade, compared with the same trader trading NYSE-listed stocks (as there is no speed upgrade to NYSE-SIP) and a trader co-located to NYSE exchanges trading Nasdaq-listed stocks (as the the traveling latency remains, i.e., data has to first travel from NYSE exchanges to Nasdaq-SIP and then back).[7] So I expect the differential impact on Nasdaq-listed stocks relative to NYSE-listed stocks, if it exists, be more pronounced Nasdaq exchanges versus other exchanges.

Then I turn to events where SIPs experienced technical glitches and the consolidated feed became corrupt or unavailable. As in all events, the technical glitch only occurred at one of the two SIPs, I am able to use the same DiD identification strategy above by constructing a matched sample of stocks whose consolidated feeds were affected and stocks that were not. Importantly, to examine SIP glitch events, one cannot rely on the consolidated feeds data such as the NYSE TAQ database because by definition they were corrupt or unavailable during the glitch. So instead I obtain direct feed data from MayStreet to compute high-frequency market liquidity measures. In addition, I exploit a unique feature of the Nasdaq-SIP glitch event on January 3, 2013 for further identification. Specifically, the glitch first occurred in even-numbered data dissemination channels ("early channels") of the Nasdaq-SIP and only occurred several minutes later in odd-numbered channels ("late channels"). So during the first glitch period, stocks allocated to the late channels can serve as an ideal control group and allows for a clean DiD analysis without the need to construct a matched sample from NYSE-listed stocks.

Findings. My major findings can be summarized as follows. First, I find that a faster consolidated feed has no statistically significant impact on market liquidity. Results from the DiD analysis show that, after the speed upgrade of Nasdaq-SIP, spreads, depth and trading volume of Nasdaq-stocks do not change *relative to* NYSE-listed stocks. Moreover, trades from retail investors do not receive better price improvements.

Second, I find an increase in low-latency trading activity, measured by cancel-to-trade ratio and strategic runs (Hasbrouck and Saar, 2013) , for Nasdaq-listed stocks *relative to* NYSE-listed

[7]See Section 2.3.2 for a detailed explanation.

stocks after the speed upgrade, In addition, the overall increase in low-latency trading activity is mainly driven by the trading in Nasdaq-listed stocks on the Nasdaq exchange, consistent with the prediction that the Nasdaq exchange is most affected by the upgrade due to its geographical proximity to the Nasdaq-SIP. In general, it suggests that fast traders have to monitor the market and update their quotes more frequently as their speed advantage from direct feeds over consolidated feeds narrows.

Third, when the SIP experiences a technical glitch and thus consolidated feeds are corrupted or unavailable, market liquidity deteriorates significantly. The DiD estimates based on a pooled sample of three glitch events show that quoted spread widens by more than 10%, depth on the NBBO drops by about 15% and trading volume falls by more than 40% for glitch-affected stocks *relative to* unaffected stocks. In sum, the findings show that the consolidated feed plays an indispensable role in today's market data infrastructure.

Contribution. My paper relates to several strands of literature. First, it is directly related to literature studying the effect of market data on trading and market quality. Brogaard and Brugler (2020) examine events where exchanges start to charge a fee for their proprietary feeds for the first time. They find that the introduction of data fees leads to a significant fall in the market volume of the fee charging exchange and it is mainly due to it having less time at the NBBO and getting less inter-market sweep (ISO) orders. Hendershott, Rysman, and Schwabe (2020) study the event when NYSE introduces its new data product, NYSE Integrated Feed, and show that there is a complementary relationship between exchange's proprietary data sales and trading activity: firms increased their share of trading on NYSE after the introduction. My paper differs from them in that I focus on the consolidated feed instead of exchanges' proprietary feeds. By examining events that exogenously affect the speed and availability of the consolidated feed, I show that the consolidated feed matters in today's market and shapes trader behavior and exchange competition, contributing to the ongoing debate over the market data reform. Ye, Yao, and Gai (2013) examine an older speed upgrade to the Nasdaq-SIP and find it has no impact on overall market liquidity. However, they do not study its impact on trader behavior and competition.

Second, my paper adds to the literature on the difference between consolidated feeds and direct feeds. O'Hara, Yao, and Ye (2014) show that odd-lot trades were missing in the consolidated feeds[8] which account for a large share of trading volume and are quite informed. They conjecture that odd-lot trader are used by the informed traders to hide their trading intention from the consolidated feeds Battalio, Corwin, and Jennings (2016) analyze a sample of high-priced stocks and show that the exclusion of odd-lot orders from consolidated feeds results in some trades being filled at worse prices. Ding, Hanna, and Hendershott (2014) compare NBBO constructed from the consolidated feeds and NBBO constructed by adding direct feeds from Nasdaq, BATS and Direct Edge exchanges. They find that the dislocation between the two NBBOs can happen quite frequently for active stocks but its duration is quite short. So the cost for low-frequent traders is small. Similarly, Bartlett and McCrary (2019) use the exchange timestamp included in the consolidated feeds to construct NBBO with zero-latency and show that profit from direct feed arbitrage is not economically significant. Hasbrouck (2019) compares the price discovery contribution of SIP compared to theoretically constructed direct feeds and finds the latter dominates at high frequency. My paper contributes to the existing studies by examining the event when the speed advantage of direct feeds over consolidated feeds largely narrows. Compared with the approach used in the existing studies that statically compares the two feeds, my paper looks at a real-life event so that it can incorporate the effect of traders' dynamic responses. However, I acknowledge that I only focus on events of two specific types, i.e., SIP's speed upgrade and glitch. So I cannot speak to other currently proposed changes to the consolidated feed such as including depth information, odd-lot quotes and auction imbalances.

Third, my paper closely relates to studies that examine the impact of trading speed on market quality[9]. Based on the adverse selection channel, theoretical models have shown that the impact of trading speed is ultimately dependent on which trader groups become faster. Market liquidity worsens if short-term informed arbitrageurs become faster (Biais, Foucault, and Moinas, 2015; Budish, Cramton, and Shim, 2015; Foucault, Hombert, and Roşu, 2016), improves if market makers become faster (Hoffmann, 2014; Jovanovic and Menkveld, 2016), and depends on market conditions (e.g., news arrival frequency and the presence of noise traders) when both trader groups become faster (Menkveld and Zoican, 2016). Empirical studies largely support the theoretical predictions

[8]Odd-lot trades were added into the consolidated feed after a SIP reform in 2013.

[9]See Menkveld (2016) for a comprehensive review on this topic.

above. For example, Brogaard, Hagströmer, Nordén, and Riordan (2015) show that market liquidity improves when market-makers opt for a speed upgrade of their co-location server to the exchange. Shkilko and Sokolov (2020) find that market liquidity improves when the microwave network between Chicago and New York is disrupted by weather conditions, which makes high-frequency arbitrageurs slower. My paper adds to the literature by examining the event of a speed upgrade to the consolidated feed, which increases the trading speed (as least the market data component) of slow traders (buy-side execution algorithms) who do not have access to direct feeds due to cost considerations or technical complexities. I find the speed upgrade has no impact on overall market liquidity, suggesting that a faster consolidated feed might have equally affected arbitrageurs and market makers. However, the arm race between them seems to intensify as low-latency trading activity increases after the upgrade. It might be due to both fast market makers and arbitrageurs responding to a faster trading speed of slow traders.

2 Data and Identification Strategy

2.1 Data

To carry out the empirical analysis, I use data from several different sources. First, I use data from the Center for Research in Security Prices (CRSP) to obtain information such as daily closing prices, market capitalization, daily trading volume and primary listing exchange, which is used in the propensity score matching exercise to be specified below. Second, I use the NYSE Trade and Quote (TAQ) database to compute a wide range of liquidity measures and trading statistics around the SIP speed upgrade events. The third dataset I use come from the SEC's Market Information Data Analytics System (MIDAS), which publishes several useful trading metrics based on direct feeds received by the SEC. For example, for each stock and exchange pair, MIDAS reports daily number and volume of new order submissions, number of cancel messages, hidden volume, and odd-lot volume. The fourth dataset is the Lobster data which is essentially an adapted copy of the Nasdaq ITCH dataset. So it contains full order book event messages such as new order submissions, cancellations and trades. The final dataset I use comes from the direct feeds collected by MayStreet[10]. Based on the direct feed messages, I am able to construct order books and compute high-frequency liquidity measures around the SIP glitch events. It is perhaps important to note that direct feeds are necessary in order to study the SIP glitch events as consolidated feed data (such as the NYSE TAQ) is either corrupt or unavailable by definition. For all events, I restrict the sample to common shares of the S&P 500 constituent stocks as low-latency traders are more active in large-cap stocks and thus more sensitive to speed upgrades and short-term data glitches. Besides, they belong to the same important index and thus are subject to similar arbitrage activities such as those between ETFs and constituent stocks.

2.2 Market liquidity and trading metrics

To assess the role of consolidated feeds, I use a wide range of liquidity and trading metrics.

Liquidity measures To assess market liquidity, I use the following common liquidity measures: relative quoted spread (*RQS*), relative effective spread (*RES*), relative realized spread (*RRS*), dollar depth at NBBO (*DollarDepth*), dollar trading volume (*Vlm*). All liquidity measures above are computed based on textbook definitions and thus formulas are omitted.

In addition to the liquidity measures defined above, I further identify retail trades following the approach in Boehmer, Jones, Zhang, and Zhang (2021). The idea is simple. It is now a common market practice for wholesalers such as Citadel and Virtu to purchase order flows from retail brokers such as Robinhood, which is known as payment for order flow. As wholesalers deem retail trades less toxic in the sense of being less informed, they are willing to provide better prices than what are available on public exchanges. The difference between the two prices are called price improvement and is mostly in the amount of sub-pennies. In contrast, sub-penny trades are not possible on public exchanges as sub-penny quotes are forbidden. So if a trade report has a sub-penny execution price, then the trade is likely to be retail and the price improvement can be easily computed as the difference between the trade price and nearest round-penny price. For example, for a trade executed at 10.011, it will be classified as a retail sell trade with a price improvement of 0.1 cent per share. After identifying retail traders, I compute share-size weighted average price improvements received by the retail trades (*PrcImp*). Importantly, I match the prices

[10]MayStreet is a US data company and supplier of SEC's MIDAS.

with the prevailing NBBOs constructed with the exchange timestamps so that I can judge whether it is a true price improvement or not. For example, a wholesaler can claim that he price improves a retail trade but actually the improvement is based on stale NBBOs from the SIP. More precisely, I only count a sub-penny trade as a price-improved trade if it executes at a price better than the prevailing NBBOs constructed without delay.[11] So it could be the case that wholesalers now have to offer more price improvements.

Trading metrics The speed upgrade might affect not only market liquidity but also the trading behavior of market participants. So I compute several trading metrics as well. The first two are the trade volume via inter-market sweep order (ISO) as a fraction of total trade volume (*ISOShr*), odd-lot trade volume as a fraction of total trade volume (*OddlotShr*). Then I compute several low-latency trading measures. The first two are the ratio of cancel order count to total trade count (*Cancel/Trade*) and the ratio of order volume of add order messages to total trade volume (*Order/Trade*). Hagströmer and Nordén (2013) show that market-making HFTs have substantially higher order-to-trade ratios than other fast traders. So an increase in the order-to-trade ratio is likely to reflect an increase in market-making related activities. However, the drawback of such quote-to-trade measures is that they also capture activities from non-HFTs such as buy-side execution algorithms. To have a better measure of low-latency HFT activities, I follow Hasbrouck and Saar (2013) and compute the number of strategic runs, which are series of linked submissions and cancellations believed to reflect the low-latency strategies of HFTs. It has been shown to have a very high correlation with true HFT activity measures, both time-series and cross-section. Moreover, following Yao and Ye (2018), I scale the raw number of runs by trading volume (*#Trade/Vlm*).

Exchange-specific metrics Except for the price improvement measure, for all other measures above, I compute them both for the whole market aggregated across all exchanges and for each exchange respectively. For example, *DollarDepth* of an exchange will be its time-weighted dollar depth at the NBBO.

[11] Anecdotal evidence shows that wholesalers might use stale SIP quotes to price their client trades. On January 13, 2017, the US SEC fined Citadel Securities $22.6 million dollars for the use of two algorithms that "did not internalize retail orders at the best price observed nor sought to obtain the best price in the marketplace." Citadel's high frequency trading strategy exploited difference between prices on SIP and the more accurate direct exchange feeds.

Table 1. Summary statistics of the matched sample for the Nasdaq-SIP upgrade event on October 24. 2016. *RQS*, *RES* and *RRS* stand for relative quoted spread, effective spread and realized spread respectively and are all in basis point. *DollarDepth* is dollar depth at NBBO in thousands. *PrcImp* is the share-size weighted average price improvements received by the retail trades, benchmarked to constructed NBBO assuming no delay. *Vlm* is dollar trading volume in millions. *OddlotShr* is odd-lot trade volume as a fraction of total trade volume. *ISOShare* is trade volume via inter-market sweep order (ISO) as a fraction of total trade volume. *Cancel/Trade* is the ratio of cancel order count to total trade count. *Order/Trade* is the ratio of order volume of add order messages to total trade volume. *#Run/Vlm* is the number of strategic runs scaled by trading volume. The sample stocks consist of 60 randomly chosen Nasdaq-listed stocks from the S&P 500 index matched with 60 NYSE-listed stocks on price, market capitalization, trading volume and industry. The sample period is from September 26 to December 1, 2016.

Variable	N	Mean	SD	Min	10%	25%	50%	75%	90%	Max
RQS	4408	4.42	3.38	1.18	1.94	2.50	3.43	5.11	7.84	31.67
RES	4408	1.59	1.39	0.40	0.72	0.89	1.21	1.74	2.75	25.69
RRS	4408	0.19	0.95	-6.87	-0.41	-0.18	0.02	0.33	0.82	15.18
DollarDepth	4408	267.46	504.98	32.50	71.83	91.26	128.89	214.86	474.68	7529.71
Vlm	4408	210.97	349.57	10.20	47.94	74.16	127.08	237.66	413.89	8827.11
PrcImp	4408	0.15	0.03	0.01	0.12	0.13	0.15	0.17	0.18	0.32
ISOShr	4408	37.01	6.17	12.93	29.16	32.80	37.04	41.18	44.76	77.96
OddlotShr	4408	11.05	6.54	0.18	3.44	6.15	9.90	15.12	20.56	35.81
Cancel/Trade	4370	23.13	9.38	6.00	13.72	16.94	21.38	27.28	34.40	122.87
Order/Trade	4370	37.35	18.18	8.89	20.84	25.96	33.54	43.76	56.46	262.08
#Run/Vlm	4325	0.60	0.33	0.08	0.29	0.39	0.53	0.73	0.96	4.75

All liquidity measures are first computed at the tick-by-tick frequency and later aggregated to a given frequency (daily for the SIP upgrade events and 30-second for the SIP glitch events). To aggregate stock variables such as *RQS* and *DepthNBBO*, I compute the time-weighted average. While for flow variables such as *RES*, *RES*, *RRS* and *RPI*, I compute dollar-volume weighted average. Note that in the computation of all liquidity measures from the TAQ data, I construct NBBOs based on exchange timestamps instead of the SIP timestamps. As illustrated above, consolidated feeds suffer processing and geographical latency, thus NBBOs constructed from SIP timestamps can be noisy and delayed compared to the real NBBOs available in the market. Table 1 reports the summary statistics of the above liquidity and trading metrics for the matched sample of Nasdaq-listed and NYSE-listed stocks around the SIP upgrade event (see below for details on the matching approach). In the appendix, I report the summary statistics by exchange for SIP upgrade event in Table A1. Summary statistics for the matched sample during the SIP glitch events can be found in Table A2, A3 and A4.

2.3 Identification strategy

2.3.1 Institutional background

Before I describe the identification strategy in detail, it helps to first introduce some institutional background regarding the market data infrastructure in the U.S. equities market. There are two so called Security Information Processors (SIPs) which disseminate quote and trade data to the general public ("consolidated feeds"). The SIP managed by NYSE ("NYSE-SIP") is responsible for securities listed on NYSE or other regional exchanges and their successors (e.g., ETPs listed on NYSE ARCA) and SIP managed by NASDAQ ("NASDAQ-SIP") is responsible for securities listed on NASDAQ. So quotes and trades of a security generated on any of the now 16 national exchanges have to be reported to the NYSE-SIP as long as it is listed on NYSE, and to the NASDAQ-SIP if it is listed on NASDAQ. In addition to the consolidated feeds published by the two SIPs, market participants can subscribe to the proprietary feeds published directly from the exchanges ("direct feeds"), which are in general faster[12] and contains more information[13] than consolidated feeds. So the fastest traders such as high-frequency trading firms (HFTs) normally use direct feeds as inputs to their trading strategies.

2.3.2 SIP speed upgrade

The major challenge with studying the impact of SIP speed upgrade is that its implementation might be endogenously driven by market conditions. For example, managing exchanges might strategically choose to upgrade their SIPs when the market is less volatile such that the message volume is lower and put less stress on the system. The unique structure of two separate SIPs in the US equities market provides a clean identification as speed upgrades to the two SIPs normally do not happen at the same time. Figure A1 plots the median processing latency of the two SIPs over the recent years and shows that NASDAQ-SIP had a significant reduction in its processing latency on October 24, 2016 while that of the NYSE-SIP barely drops. NYSE-SIP had a reduction in its processing latency at the end of 2018 and again in July 2020 while that of the NASDAQ-SIP

[12]The reason why SIP is slow is not only because of its less technologically advanced hardware and software, but also its current structure. A common example used is the case where a trader physically based at Carteret trades a Tape A/B security. A quote update on Nasdaq will have to be sent to NYSE's SIP at Mahwah, processed by it, and then sent back to the trader. However, if the trader subscribes to Nasdaq's direct feed, he will receive the update without such traveling latency.

[13]For example, direct feed contains depth information, add-lot quotes and auction imbalance information. than the consolidated feed.

Figure 1. SIP speed upgrades. Starting from August 2015, each TAQ message has two timestamps: exchange timestamp (when the message is registered at the exchange from which it originates) and SIP timestamp (when the message is disseminated by the SIP). Thus I can compute the SIP latency by the difference between the two timestamps. To compute the processing latency of the NYSE-SIP, I use quotes in General Electric originating from the NYSE exchange so that there is minimal traveling latency. By the same token, to compute the processing latency of the Nasdaq-SIP, I use quotes in Apple originating from the Nasdaq exchange. The figures plots the daily median latency for NYSE-SIP and Nasdaq-SIP respectively.

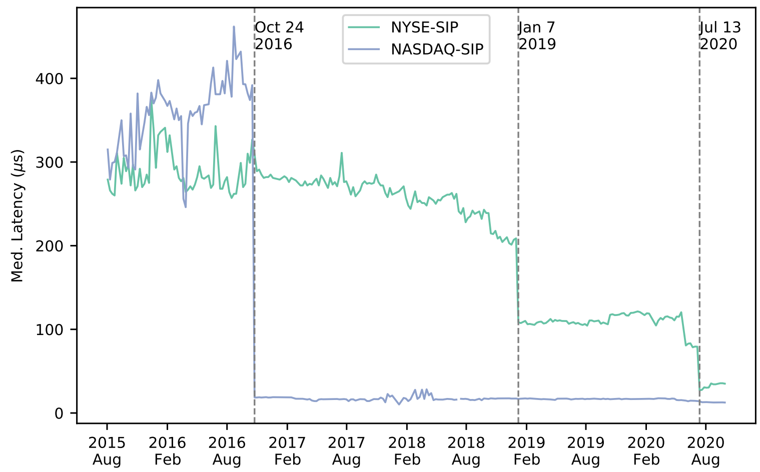


Table 2. Propensity score matching: SIP speed upgrade. This table reports results from the propensity score matching for the SIP speed upgrade. The treatment group consists of 60 randomly chosen NASDAQ-listed stocks are matched with 60 NYSE-listed stocks on price, trading volume, market capitalization and Fama and French 12 industry classification. I use one-to-one nearest neighbor propensity score matching (PSM), without replacement.

Variable	Sample	N	Mean	SD	10%	25%	50%	75%	90%
Price	Control	60	92.19	78.49	26.62	46.32	70.47	118.16	177.15
	Treatment	60	91.44	113.88	28.83	39.57	64.74	90.74	119.90
MarketCap	Control	60	37.39	48.35	8.90	12.77	23.61	33.12	66.91
	Treatment	60	38.31	54.77	7.01	11.33	19.08	40.98	89.70
DollarVolume	Control	60	224.68	203.55	60.38	95.47	143.72	263.84	503.85
	Treatment	60	265.42	327.89	49.12	79.16	166.70	272.31	579.30
PSM Score	Control	60	0.35	0.16	0.15	0.27	0.32	0.54	0.57
	Treatment	60	0.36	0.17	0.15	0.27	0.32	0.54	0.59

stayed the same.

In the following analysis, I zoom in onto the first event when the NASDAQ-SIP had the speed upgrade as the magnitude of its processing latency reduction is most significant: the median processing latency of NASDAQ-SIP is around 350 microseconds before the speed upgrade and drops more than 10 times to around 20 microseconds after. Given the difference in processing-latency reduction between the two SIPs, it allows for a different-in-difference (DiD) analysis based on a matched sample of Nasdaq-listed and NYSE-listed stocks. Specifically, I first randomly chose 60 Nasdaq-listed stocks from the S&P 500 index and then match them with 60 NYSE-listed stocks from the same index on price, trading volume, and market capitalization.[14] In addition, as Nasdaq-listed and NYSE-listed stocks are not evenly distributed across industries. I add Fama and French 12 industry classification as a further matching variable following Brogaard, Ringgenberg, and Rösch (2020).[15] Moreover, I use one-to-one nearest neighbor propensity score matching (PSM), without replacement. Table 2 reports the matching results and shows that matching is successful as all matching variables of of the two samples have similar support. The sample period covers from September 26, 2016 to December 1, 2016, which roughly spans a one-month window before and after the event date. A two-month length is chosen to strike a balance between a too-long window which might include other events and a too-short window which might not generate sufficient statistical power.

The DiD identification approach requires the standard parallel trends assumption, which means the treatment group would have evolved in a similar fashion to the control group if there was no speed upgrade to the SIP. Figure A1 in the appendix plots time-series of the liquidity and trading metrics around the upgrade event and visual evidence suggests that it is indeed the case. In other words, the parallel trends assumption is supported.

Nasdaq-listed vs. NYSE-listed stocks Using the matched sample approach, I estimate the difference-in-difference (DiD) regressions as follows:

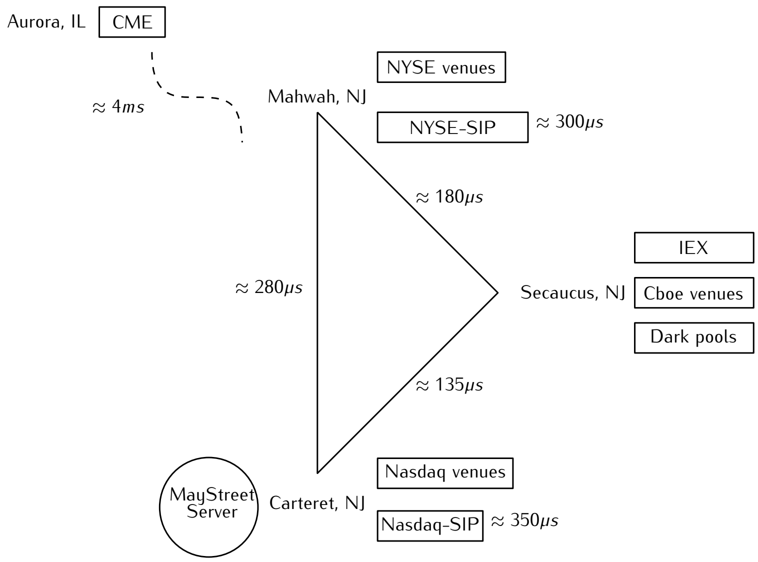
$$metric_{i,t} = \alpha_i + \beta After_{i,t} + \gamma After_{i,t} \times NasdaqStock_{i,t} + \epsilon_{i,t} \tag{1}$$

where $metrc_{i,t}$ is the liquidity or trading metric of stock i on day t. α_i is the stock fixed effects. $NasdaqStock_{i,t}$ is a dummy variable that equals one if stock i is a Nasdaq-listed stock, and equals zero otherwise. $After_{i,t}$ is dummy variable that equals one on and after October 24, 2016, and equals zero otherwise. Standard errors are clustered at the stock level.

[14]There are 107 Nasdaq-listed stocks in the S&P 500 index and 327 NYSE-listed stocks. So the number of candidate NYSE-stocks is roughly five times that of the treated Nasdaq stocks. Such a ratio makes a proper matching possible.

[15]Perhaps it is worth noting that they use Fama and French 48 industry classification. The reason for me to use the simpler version is due to the relatively small size of the sample. Not every industry has stocks from both tapes in my sample, which makes the logit regression not converge.

Figure 2. The New Jersey "Equity Triangle". This figure sketches the geographical locations of major US equity exchanges and two SIPs. Besides, it shows the estimated traveling latencies between three data centers and processing latencies at the two SIPs before the upgrade event on October 24, 2016.[16]



Nasdaq venue vs. other venues The second identification strategy utilizes the unique geographical locations of exchanges and two SIPs. As illustrated in Figure 2, all major US equity exchanges and the two SIPs are located in a triangular area in New Jersey, with Nasdaq venues (Nasdaq, BX, and PSX) and the Nasdaq-SIP at Carteret, NYSE venues (NYSE, Arca, American) and the NYSE-SIP in Mahwah and Cboe venues (BZX, BYX, EDGX and EDGA) in Secaucus. Thus the speed upgrade of the Nasdas-SIP will potentially have differential impact on the trading on exchanges located at different locations.

The idea is as follows: Note that for a trader who subscribes to SIP feed, the total latency of receiving a message from one exchange consists of two parts: processing latency of the message at the SIP and traveling latency of the message, first from the exchange to SIP and then from SIP to the trader. As an example, consider trading in Nasdaq-listed stocks (thus messages from all exchanges have to be reported to and processed at Nasdaq-SIP). For a trader who co-locates to Nasdaq venues ("Trader A"), the total latency of SIP is dominated by the processing latency as there is barely no traveling latency from Nasdaq venues to Nasdaq-SIP and from Nasdaq-SIP to the trader as they are all located at the same geographical location. In contrast, for a trader who co-locates to NYSE venues ("Trader B"), a message from NYSE venues (at Mahwah) has to first travel to Nasdaq-SIP (at Carteret), be processed there, and then travels back from Nasdaq-SIP (at Carteret) to the trader (at Mahwah). So after the speed upgrade which reduces the processing latency of the Nasdaq-SIP from 350 microseconds to less than 20 microseconds, the total SIP latency for Trader A is only about 20 microseconds. While for Trader B, the total latency will be 20 microseconds of processing latency plus a round trip traveling latency between Carteret and Mahwah, which is roughly 560 microseconds through optical fiber and not affected by the speed upgrade. In other words, the advantage of subscribing to direct feeds over SIP feeds will largely diminishes for Trader A while stays significant for Trader B.

Utilizing the geographical latency explained above, I use the following regression specification to study the impact of the SIP speed upgrade on market liquidity and trading behaviors:

$$\begin{aligned} metric_{i,e,t} = \alpha_{i,e} + \beta_t + \gamma_1 After_{i,e,t} \times NasdaqStock_{i,e,t} + \gamma_2 After_{i,e,t} \times NasdaqVenue_{i,e,t} \\ + \gamma_3 After_{i,e,t} \times NasdaqStock_{i,e,t} \times NasdaqVenue + \epsilon_{i,e,t} \end{aligned} \tag{2}$$

where $metric_{i,e,t}$ is the liquidity or trading metric for stock i, traded on exchange e and on day t. $\alpha_{i,e}$ controls for the stock-venue fixed effects, which is is potentially important as a stock can trade at a fixed pattern on a particular exchange. For example, although Nasdaq-listed stocks can trade at any of the 16 exchanges, auctions are only held at Nasdaq, the listing exchange. Besides, Nasdaq offers a different fee for stocks listed on its venue compared to stocks listed on NYSE. β_t controls for the day-fixed effects which impacts all stock-exchange pairs on the date.[17] $After_{i,t}$ is dummy variable that equals one after Oct 24, 2016, and equals zero otherwise. $NasdaqStock_{i,e,t}$ is a dummy variable that equals one if stock i is a Nasdaq-listed stock, and equals zero otherwise. $NasdaqVenue_{i,e,t}$ is a dummy variable that equals one if the metric is computed based on exchange e, and equals zero otherwise.[18] So the coefficient γ_3 captures the triple-difference-in-difference effect, that is, the change in the cross-venue difference for Nasdaq-listed stocks versus NYSE-listed stocks after the SIP speed upgrade. Standard errors are clustered at the stock level.

Note that for the DiD identification above to be valid, one needs the parallel assumption to hold on the exchange level, that is, liquidity and trading metrics would have evolved similarly on the Nasdaq exchange relative to other exchanges if there was no speed upgrade to the Nasdaq-SIP.

[17] Note that other interactions terms are absorbed due to the inclusion of stock-venue and day-fixed effects.

[18] I only include four exchanges NYSE-Arca, Nasdaq, BZX and EDGX as they all adopt a maker-taker model. Moreover, NYSE is excluded as it has only started to trade Nasdaq-listed stocks since April 2018.

Table 3. Recent major SIP glitches. This table lists some key information about the three major SIP glitch events that happened in recent years.

Date	Time	Duration	SIP	Trading halt
January 3, 2013	13:33 - 13:51	18 minutes	Nasdaq-SIP	No[a]
October 30, 2014	13:07 - 13:34	27 minutes[b]	NYSE-SIP	No
August 12, 2019	15:15 - 15:27	12 minutes[b]	NYSE-SIP	No

[a] There is no market-wide trading halt. EDGX and EDGA halted trading for Nasdaq-listed stocks after 13:42.

[b] In both events, The NYSE shifted operations to its disaster recovery site in Chicago after the glitch was solved.

Figure A2 in the appendix plots the time-series of the liquidity and trading metrics series on several exchanges and visual evidence supports the parallel trends assumption.

2.3.3 SIP glitches

Pooling SIP glitch events The identification strategy for studying the impact of SIP glitches is the same as above and uses the unique structure of two SIPs in the US equities market. Table 3 lists major SIP glitches in recent years.[19] For each event, it is either Nasdaq-SIP or NYSE-SIP that suffered a glitch, not both. So when NYSE-SIP experienced a glitch, I expect trading in NYSE-listed stocks to be more affected than that in Nasdaq-listed stocks and vice versa when Nasdaq-SIP experienced a glitch. Importantly, as Figure A3, A4 and A5 in the appendix show, during all three SIP glitch events, direct feeds were largely unaffected. So the identified effect is not due to trade disruption as in the Nasdaq "Flash Freeze" event on August 22, 2013.

For each event, I use the same matching approach as above to construct a matched sample of 60 treated stocks (e.g., Nasdaq-listed stocks when Nasdaq-SIP suffers a glitch) and 60 control stocks (e.g., NYSE-listed stocks when Nasdaq-SIP suffers a glitch). Table 4 reports the matching results and shows it is quite successful. The sample period for each event starts from 30 minutes before the glitch starts and to the glitch ends. Then I run a standard DiD regression as in Equation 3 below:

$$metric_{i,d,t} = \alpha_{i,d} + \beta After_{i,d,t} + \gamma After_{d,t} \times Treated_{i,d} + \epsilon_{i,d,t}. \tag{3}$$

[19] Another well-known SIP glitch is the Nasdaq "flash freeze" on August 22, 2013, which resulted in a 3-hour trading halts of Nasdaq-listed stocks. However, the event is excluded as one lacks the treatment group to study the impact.

Table 4. Propensity score matching: SIP glitch events This table reports results from the propensity score matching for the three SIP glitch events. For each event, the treatment group consists of 60 randomly chosen stocks whose consolidated feeds are affected by the SIP glitch and are matched with 60 unaffected stocks on price, trading volume, market capitalization and Fama and French 12 industry classification. I use one-to-one nearest neighbor propensity score matching (PSM), without replacement.

Variable	Sample	N	Mean	SD	Min	10%	25%	50%	75%	90%	Max
Price	Control	180	83.67	80.04	2.53	18.49	36.01	58.72	100.92	167.08	510.27
	Treatment	180	92.99	115.83	6.57	21.93	37.24	66.74	105.09	164.26	1123.04
MarketCap	Control	180	308.48	845.51	18.01	52.57	81.03	143.40	266.96	568.07	10404.77
	Treatment	180	334.65	519.98	23.74	55.09	90.04	151.50	349.03	616.01	3436.72
DollarVolume	Control	180	96.80	232.23	1.62	14.04	25.00	45.36	88.05	186.02	2839.34
	Treatment	180	108.49	165.48	8.42	17.37	26.98	46.53	125.33	251.02	1450.20
PSM Score	Control	180	0.62	0.23	0.17	0.23	0.48	0.69	0.80	0.84	0.98
	Treatment	180	0.65	0.25	0.17	0.22	0.47	0.72	0.84	0.93	1.00

where $metrc_{i,d,t}$ is the liquidity or trading metric of stock i on event day d during the 30-second time interval t. $\alpha_{i,d}$ is the stock-day fixed effect. $Treated_{i,d}$ is a dummy variable that equals one if stock i is a Nasdaq-listed stock on January 3, 2013 and NYSE-listed stock on October 30, 2014 and August 12, 2019. $After_{d,t}$ is a dummy variable that equals one after the glitch starts, and equals zero otherwise. Standard errors are clustered at the stock-day level. It is perhaps worth noting that the three events happened at different intraday periods and have glitches both at the Nasdaq-SIP and NYSE-SIP, thus pooling all three events in the regression helps alleviate some concerns from a possibly imperfect matching.

Zoom in onto the event on January 3, 2013 I exploit a unique feature of the Nasdaq-SIP glitch event on January 3, 2013 for a further identification. Specifically, there are six channels through which the Nasdaq-SIP disseminates its data feeds and Nasdaq-listed stocks are allocated into the six channels *alphabetically*.[20] More importantly, as shown in Table 5, the Nasdaqp-SIP glitch on January 3, 2013 first occurred at three even-numbered channels and occurred only a few minutes later at the other three odd-numbered channels. Thus stocks in the "late" three channels can serve as a ideal control group during the time period between the start of the first and second glitch.

[20] **Ye2012** use such feature to study the potential quote stuffing behavior of HFTs and find that messages of stocks within the same channels are more correlated. The allocation of symbols in each of the six channels are according to the alphabetical order. Although the allocation rule might be such that the total message volume in each channel show be more or less the same so that no channel will be suffering constantly high message volume and having higher latency

Table 5. Channel assignment of NASDAQ-SIP and outage order. This table shows the symbol allocation across the six data dessimination channels of the Nasdaq-SIP. Moreover, it shows the starting and ending time of the glitch for trades and quotes in each channel.

Outage order	Channel	Quote outage period	Trade outage period
"Late"channels	Channel 1 (Symbols A-CDZ) Channel 3 (Symbols FE-LKZ) Channel 5 (Symbols PC-SPZ)	13:37:22 - 13:48:19	13:36:51 - 13:51:14
"Early" channels	Channel 2 (Symbols CE-FDZ) Channel 4 (Symbols LL-PBZ) Channel 6 (Symbols SQ-ZZZ)	13:33:11 - 13:48:21	13:33:11 - 13:51:15

So I use two difference-in-difference (DiD) regressions to study the unique event on January 3, 2013. The first regression Equation 4 focuses on the period when the early channels experienced data outage while not yet for the late channels.

$$metric_{i,t} = \alpha_i + \beta Period1_{i,t} + \gamma_1 Period1_{i,t} \times EarlyChannel_{i,t} + \gamma_2 Period1_{i,t} \times LateChannel_{i,t} + \epsilon_{i,t} \quad (4)$$

where $metrc_{i,t}$ is the liquidity or trading metric of stock i in time interval t. α_i is the stock fixed effects. $Period1_{i,t}$ is dummy variable that equals one between the start of the glitch at early channels and late channels, that is, between 13:33:11 and 13:36:51, and equals zero otherwise. Standard errors are clustered at the stock level. $EarlyChannel_{i,t}$ is dummy variable that equals one if stock i belongs to one of the early channels, and equals zero otherwise. $LateChannel_{i,t}$ is a dummy variable that equals one if stock i belongs to one of the late channels, and equals zero otherwise. Note for the first regression, I include stocks from all early, late and normal channels. The sample period is between 30 minutes before the glitch started at early channels and the start of glitch at late channels. So we should expect γ_1 to be significant while γ_2 insignificant if the first glitch only affects early channels.

The second regression focuses on the period when the late channels started to experience the glitch and is specified in Equation 5 as below.

$$metric_{i,t} = \alpha_i + \beta Period2_{i,t} + \gamma Period2_{i,t} \times LateChannel_{i,t} + \epsilon_{i,t} \quad (5)$$

where $metrc_{i,t}$ is the liquidity or trading metric of stock i in time interval t. α_i is the stock fixed effects. $Period2_{i,t}$ is dummy variable that equals one between the start of the glitch at late channels and the end of glitch, that is, between 13:36:51 and 13:51:15, and equals zero otherwise. Standard errors are clustered at the stock level. $LateChannel_{i,t}$ is dummy variable that equals one if stock i belongs to one of the late channels, and equals zero otherwise. Note that for the second regression, I only include stocks of the late channels and normal channels as stocks of the early channels have already been "treated" and thus violates the parallel trends condition. The sample period is between the start of the glitch at early channels and the end of all glitches.

3 Results

I next take the identification strategies developed in the previous section to the data and examine the role of consolidated feeds in the current US equities market. I first look at the event of a speed upgrade to the consolidated feeds and then turn to events where the consolidated feeds experience a technical glitch and thus become corrupted or unavailable to market participants.

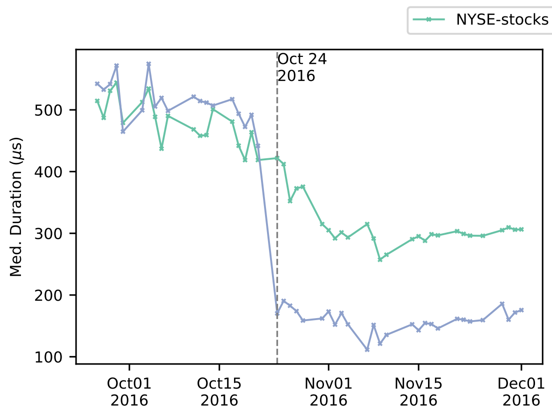
3.1 SIP speed upgrade

NBBO dislocations Before turning to liquidity and trading, it is worthwhile to see to what extend the SIP speed upgrade make SIP prices more reliable. Following Ding, Hanna, and Hendershott (2014) and Bartlett and McCrary (2019), I construct two versions of NBBO, one based on the SIP-timestamp (SIP-NBBO) and the other based on the Participant-timestamp[21] (Direct-NBBO). Then dislocations between the two NBBOs occur when the SIP-NBBO diverges from the prevailing Direct-NBBO. Moreover, by the same token, I construct the best bid and offer prices on the Nasdaq-exchange based on the SIP-timestamp (SIP-Nasdaq-BBO) and Participant-timestamp (Direct-Nasdaq-BBO). After identifying the dislocations, I compute both the number of dislocations and the median dislocation duration for each stock-day pair in the matched sample. Figure 3 plots the cross-section average of the two measures for both the NBBO (3a) and Nasdaq-BBO (3b). As expected, Nasdaq-stocks see a significant decrease of about 300 microseconds in the median dislocation duration for both the NBBO and Nasdaq-BBO after the speed upgrade. While there is

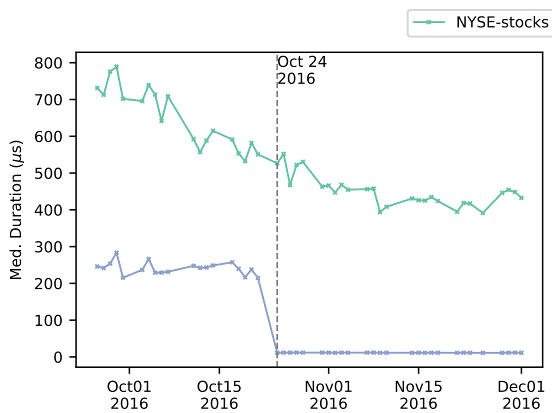
[21]Participant times records when a quote update or trade is registered at the exchange's matching engine and thus suffers no processing and traveling latency at the SIP.

Figure 3. SIP speed upgrade and NBBO dislocations. Following Ding, Hanna, and Hendershott (2014) and Bartlett and McCrary (2019), I construct two versions of NBBOs, one based on the SIP timestamps (SIP-NBBO) and the other based on the Participant timestamps (Direct-NBBO) and then identify the dislocations between the two NBBOs. By the same token, I construct two BBOs on the Nasdaq exchange, one based again on SIP timestamps and the other on Participant timestamps.

(a) This figure plots the median NBBO dislocation duration and number of NBBO dislocations for NYSE-Stocks and Nasdaq-stocks respectively. The time series is computed as cross-section average.



(b) This figure plots the median Nasdaq-BBO dislocation duration and number of Nasdaq-BBO dislocations for NYSE-Stocks and Nasdaq-stocks respectively. The time series is computed as cross-section average.



no structural shifts in the the number of dislocations as the SIP processing latency is not completely eliminated after the speed upgrade and the geographical latency remains[22], it drops for Nasdaq-stocks *relative to* NYSE-stocks. The difference is more pronounced for Nasdaq-BBOs as there is no geographical latency. To provide some numbers, the number of dislocations in the NBBO (Nasdaq-BBO) for Nasdaq-stocks drops by more than 746 (1854) than NYSE-stocks, or roughly 15% (41%) relative to pre-upgrade level.

[22]For example, quotes in Nasdaq-stocks from NYSE and Cboe venues incur extra traveling latency to reach Nasdaq-SIP.

Table 6. Difference-in-difference regression: market-wide impact of the Nasdaq-SIP upgrade. This table shows results from the following difference-in-difference regression:

$$metric_{i,t} = \alpha_i + \beta After_{i,t} + \gamma After_{i,t} \times NasdaqStock_{i,t} + \epsilon_{i,t}$$

where $metrc_{i,t}$ is the liquidity or trading metric of stock i on day t. α_i is the stock fixed effects. $Treated_{i,t}$ is dummy variable that equals one if stock i is a Nasdaq-listed stock, and equals zero otherwise. $After_{i,t}$ is dummy variable that equals one after Oct 24, 2016, and equals zero otherwise. Standard errors are clustered at the stock level. The sample stocks consist of 60 randomly chosen Nasdaq-listed stocks from the S&P 500 index matched with 60 NYSE-listed stocks on price, market capitalization, trading volume and industry. The sample period is from September 26 to December 1, 2016.

(a) Liquidity metrics. *RQS*, *RES* and *RRS* stand for relative quoted spread, effective spread, realized spread and price impact respectively and are all in basis point. *DollarDepth* is dollar depth at NBBO in thousands. *Vlm* is dollar trading volume in millions. *PrcImp* is share-size weighted average price improvements received by the retail trades, benchmarked to direct-NBBOs.

	RQS	RES	RRS	DollarDepth	Vlm	PrcImp
After	0.53***	0.23***	0.00	−26.27*	64.09***	0.00
	(0.09)	(0.06)	(0.05)	(15.07)	(12.44)	(0.00)
After x NasdaqStock	0.06	−0.10	−0.07	7.10	3.59	0.00
	(0.14)	(0.08)	(0.07)	(18.61)	(30.33)	(0.00)
R^2 (%)	7.67	1.14	0.10	0.78	2.84	0.12
N	4408	4408	4408	4408	4408	4408
Stock F.E.	Yes	Yes	Yes	Yes	Yes	Yes

(b) Trading metrics. *ISOShr* is trade volume via inter-market sweep order (ISO) as a fraction of total trade volume. *OddlotShr* is odd-lot trade volume as a fraction of total trade volume. *Cancel/Trade* is the ratio of cancel order count to total trade count. *Order/Trade* is the ratio of order volume of add order messages to total trade volume. *#Run/Vlm* is the number of strategic runs scaled by trading volume.

	ISOShr	OddlotShr	Cancel/Trade	Order/Trade	#Run/Vlm
After	−1.89***	−0.52**	−6.39***	−10.20***	−0.12***
	(0.33)	(0.21)	(0.84)	(1.60)	(0.03)
After x NasdaqStock	1.26**	0.00	3.79***	5.71***	0.08**
	(0.50)	(0.31)	(0.98)	(1.79)	(0.04)
R^2 (%)	2.01	1.04	12.16	9.40	2.81
N	4408	4408	4370	4370	4325
Stock F.E.	Yes	Yes	Yes	Yes	Yes

Market-wide impact I first examine whether the speed upgrade to Nasdaq-SIP affects market-wide liquidity by running the DiD regression specified in Equation 1. Table 6a reports the results. It shows that the coefficient on the interaction term *After* × *NasdaqStock*, which captures the DiD effect, is statistically insignificant for all liquidity measures, suggesting that in general a faster consolidated feed has no impact on market-wide liquidity.

Although the speed upgrade of Nasdaq-SIP does not affect market-wide liquidity, it has a significant impact on trader behavior as shown by results in Table 6b. First, the coefficient on the interaction term *After* × *NasdaqStock* is positive and significant for *ISOShr*, suggesting that traders use more ISO orders in the trading of Nasdaq-listed stocks than NYSE-listed stocks. As Chakravarty, Jain, Upson, and Wood (2012) shows, ISO orders are mainly used by informed institutional traders to sweep through top-of-book depth across all exchanges simultaneously. A more up-to-date and reliable consolidated feed for Nasdaq-listed stocks might explain why traders submit more ISO orders after the upgrade. However, the DiD effect of 1.26% is quite small in both absolute and relative magnitude. Second, the coefficients on all three low-latency measures *Cancel/Trade*, *Add/Trade* and *#Run/Vlm*, are positive and significant. As such measures are viewed as proxies for algorithmic trading (Hendershott, Jones, and Menkveld, 2011) and *#Run/Vlm* especially characterizes HFT activities, it indicates that there is an increase in low-latency trading activities for Nasdaq-listed stocks *relative to* NYSE-listed stocks after the Nasdaq-SIP upgrade. It might be due to the fact that HFTs will have to monitor the market and update their quotes more frequently given that the speed advantage of direct feeds over consolidated feeds greatly vanishes.

Exchange-specific impact After examining the impact of the Nasdaq-SIP speed upgrade on *market-wide* liquidity and trading, now I look at whether the impact might be different on the *Nasdaq exchange versus other exchanges* due to their geographical locations illustrated in Section 2. So I run the triple-DiD regression specified in Equation 2 with exchange-specific metrics as dependent variables. Table 7 reports the regression results. Again, it shows that the Nasdaq-SIP speed upgrade has largely no impact on liquidity as the coefficient on the triple interaction term *After* × *NasdaqStock* × *NasdaqVenue* is insignificant for all liquidity measures.

Turning to the regression results regarding trading metrics, coefficient on the triple interaction term *After* × *NasdaqStock* × *NasdaqVenue* is significant for ISO share and two low-latency metrics and has the same sign as in the DiD results above. In contrast, coefficient on the interaction term *After* × *NasdaqStock* is insignificant, suggesting that the impact of the Nasdaq-SIP speed upgrade is mainly driven by changes in the trading of Nasdaq-listed stocks on the Nasdaq exchange, consistent with the conjecture illustrated in Section 2. Interestingly, for all trading metrics, the coefficient sign on the interaction term *After* × *NasdaqVenue* is opposite to that on the triple interactions term *After* × *NasdaqStock* × *NasdaqVenue* and the magnitude of the latter is larger. In other words, the

Table 7. Triple-difference-in-difference regression: exchange-specific impact of SIP upgrade. This table shows results from the triple-difference-in-difference regression

$$metric_{i,e,t} = \alpha_{i,e}+\beta_t+\gamma_1 After_{i,e,t}\times NasdaqStock_{i,e,t}+\gamma_2 After_{i,e,t}\times NasdaqVenue_{i,e,t}+\gamma_3 After_{i,e,t}\times NasdaqStock_{i,e,t}\times NasdaqVenue+\epsilon_{i,e,t}$$

where $metric_{i,e,t}$ is the liquidity or trading metric for stock i, traded on exchange e and on day t. $\alpha_{i,e}$ controls for the stock-venue fixed effects and β_t controls for the day-fixed effects. $After_{i,t}$ is dummy variable that equals one after Oct 24, 2016, and equals zero otherwise. $NasdaqStock_{i,e,t}$ is dummy variable that equals one if stock i is a Nasdaq-listed stock, and equals zero otherwise. $NasdaqVenue_{i,e,t}$ is dummy variable that equals one if the metric is computed based on exchange e, and equals zero otherwise. Standard errors are clustered at the stock level.The sample stocks consist of 60 randomly chosen Nasdaq-listed stocks from the S&P 500 index matched with 60 NYSE-listed stocks on price, market capitalization, trading volume and industry. The sample period is from September 26 to December 1, 2016.

(a) Liquidity metrics. All measures have been defined above. But note that now they are measures specific to an exchange. For example, *DollarDepth* of an exchange will be its time-weighted dollar depth at the NBBO.

	RQS	RES	RRS	DollarDepth	Vlm
After	0.77	0.14***	−0.11***	−2.91	4.47***
	(0.61)	(0.03)	(0.02)	(1.88)	(0.87)
After x NasdaqStock	0.87	−0.01	−0.01	0.96	1.77
	(0.68)	(0.04)	(0.04)	(2.23)	(2.37)
After x NasdaqVenue	−0.09	0.00	0.00	0.52	5.14***
	(0.56)	(0.01)	(0.01)	(0.98)	(0.86)
After x NasdaqStock x NasdaqVenue	−0.95	0.02	−0.02	−0.18	6.08
	(0.62)	(0.01)	(0.02)	(1.63)	(3.77)
R^2 (%)	0.36	3.43	0.95	0.28	3.00
N	17632	17632	17632	17632	17632
Stock-Venue F.E.	Yes	Yes	Yes	Yes	Yes

(b) Trading metrics. All measures have been defined above.

	ISOShr	OddlotShr	Cancel/Trade	Order/Trade
After	−2.45***	−0.44*	−6.14***	−8.43***
	(0.44)	(0.25)	(0.84)	(1.08)
After x NasdaqStock	1.00*	0.12	2.45**	2.60*
	(0.56)	(0.33)	(1.13)	(1.46)
After x NasdaqVenue	−0.35	0.03	−2.03***	−3.93***
	(0.24)	(0.14)	(0.58)	(0.91)
After x NasdaqStock x NasdaqVenue	0.67*	−0.66***	3.68***	6.17***
	(0.35)	(0.21)	(0.79)	(1.18)
R^2 (%)	2.89	0.53	5.84	6.43
N	17632	17632	17480	17480
Stock-Venue F.E.	Yes	Yes	Yes	Yes

speed upgrade has an opposite impact on Nasdaq-listed stocks versus NYSE-listed stocks when they trade on the Nasdaq exchange. The results above are consistent with the conjecture that the Nasdaq-SIP speed upgrade has a larger effect on Nasdaq-stocks traded on the Nasdaq exchange given the geographical proximity of the two.

In summary, I do not find evidence that a faster consolidated feed improves or worsens market liquidity in general. However, the speed upgrade leads to an increase in low-latency trading, especially for Nasdaq-stocks traded on the Nasdaq-exchange. Past empirical studies show that an exogenous shock to a subset of low-latency traders affects market liquidity. For example, market liquidity improves when high-frequency market makers become faster (Brogaard, Hagströmer, Nordén, and Riordan (2015)) or high-frequency arbitrageurs become slower (Shkilko and Sokolov (2020)). The result that a faster consolidated feed has no impact on overall market liquidity might be due to it affecting both types of low-latency traders equally. However, both trader groups seem to step up their trading activity when slow traders become faster. The speed upgrade of Nasdaq seems to have created a wasteful arm race between slow traders and fast traders, resulting in no good for the market quality.

3.2 SIP glitches

The results so far show the impact of a speed upgrade to consolidated feeds, now I turn to SIP-glitch events and look at what happens to the market when consolidated feeds are not available.

3.2.1 Pooling all three events

I first run the DiD regression specified in Equation 3 where I pool all three major SIP glitch events as detailed in Table 3. Table 8 reports the regression results. It shows that in general, treatment stocks (i.e., stocks whose consolidated feeds are corrupted or unavailable due to the SIP glitch) see their market liquidity worsens by all common measures. Specifically, coefficients on the interaction dummy *After* $\times$ *Treated*, which capture DiD effects, show that quoted spread and effective spread of treatment stocks increase by 0.46 and 0.11 basis points more than control stocks respectively, which correspond to about 10.7% (0.46 / 4.29) and 7.7% (0.11 / 1.43) relative to their unconditional means[23]. Out of all liquidity metrics, trading volume is most affected, with trading volume in treatment stocks falling by nearly 40% (-41.39 / 108.84) more than control stocks. In addition, the coefficient on the *After* dummy is significantly positive for trading volume, suggesting that traders do not simply refrain from trading all together, but to some extent shift their trading volume from

[23]Unconditional means are simply the average across all stock-days for the pre-event period.

Table 8. Difference-in-difference regression: pooling all three SIP-glitch events. This table show regression results from the following difference-in-difference regression:

$$metric_{i,d,t} = \alpha_{i,d} + \beta After_{i,d,t} + \gamma After_{d,t} \times Treated_{i,d} + \epsilon_{i,d,t}.$$

where $metrc_{i,d,t}$ is the liquidity or trading metric of stock i on event day d during the 30-second time interval t. $\alpha_{i,d}$ is the stock-day fixed effect. $Treated_{i,d}$ is a dummy variable that equals one if stock i is a Nasdaq-listed stock on January 3, 2013 and NYSE-listed stock on October 30, 2014 and August 12, 2019. $After_{d,t}$ is a dummy variable that equals one after the glitch starts, and equals zero otherwise. Standard errors are clustered at the stock-day level. The sample stocks consist of 180 treated stocks (i.e., affected by the SIP glitch) matched with 180 control stocks on price, market capitalization, trading volume and industry. The sample period is between 30 minutes before the start of the SIP glitch and the end of it. *Depth5lvl*

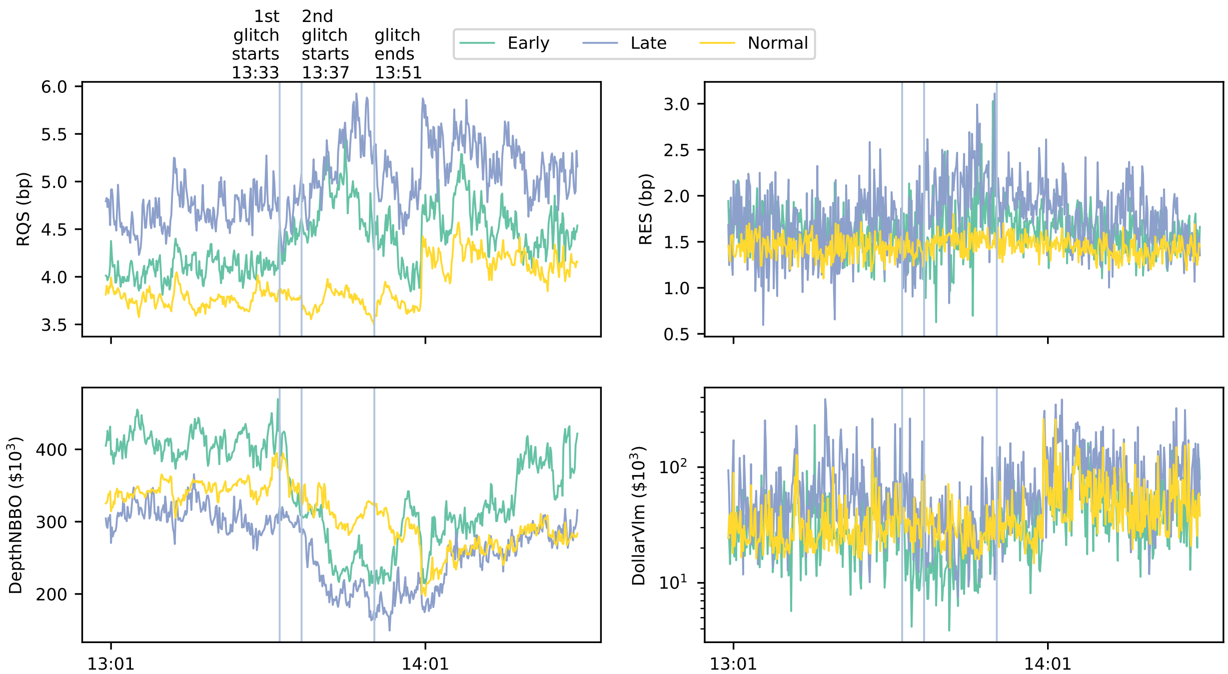
	RQS	RES	RRS	DollarVlm	DepthNBBO	Depth5lvl
After	0.04	−0.02	−0.01	26.72***	−16.70	−124.56***
	(0.07)	(0.03)	(0.07)	(9.39)	(11.54)	(36.17)
After x Treated	0.46***	0.11**	0.51***	−41.39***	−34.83**	−143.40**
	(0.12)	(0.04)	(0.09)	(11.18)	(14.88)	(58.98)
R^2 (%)	1.72	0.12	0.53	0.18	1.93	6.10
N	34964	30497	30497	34964	34964	34964
Stock-Event F.E.	Yes	Yes	Yes	Yes	Yes	Yes

treatment stocks to control stocks. As for the two depth measures, they follow a similar pattern. For example, depth at the NBBO for treatment stocks drops by about 15% (-34.83 / 240.03) more than control stocks. Note that coefficients on the dummy term *After* are negative for depth within five levels from NBBO as well, indicating that depth of control stocks decrease as well. The result is perhaps not surprising as trading in the two matched sample of stocks are correlated, either due to them being in the same market index or in the same industry. So when treatment stocks become illiquid, the illiquidity can quickly spread to the control stocks. Such illiquidity contagion through informationally correlated assets are well modeled in Cespa and Foucault (2014).

3.2.2 Zooming in onto the event on January 3, 2013

Then I zoom in onto the SIP glitch event on January 3, 2013. First, Figure 4 plots several liquidity measures around the two glitches and visually show that stocks in the early channel and late channel are affected sequentially. Then I run the DiD regressions specified in Equation 4 and 5 respectively, formally testing the DiD effect. Table 9 reports the results. The first DiD regression is run over the first period of the glitch, that is, when stocks of the even-numbered channels started to experience a glitch but not yet for stocks in the odd-numbered channels. The coefficient on the

Figure 4. Liquidity metrics around the Nasdaq-SIP glitch event on January 3, 2013. This figure plots several liquidity metrics for three groups of stocks respectively. The "Early" group includes Nasdaq-listed stocks belonging to the data channels which were first hit by the glitch. The "Late" groups includes Nasdaq-listed stocks belonging to the data channels which were hit by the glitch later. The "Normal" group includes a matched sample of NYSE-listed stocks. *RQS*, *RES* stand for relative quoted spread and relative effective spread in basis point. *DollarVlm* is dollar trading volume in millions. *DepthNBBO* is dollar depth at NBBO in thousands. The first two vertical lines indicate the start of the glitch at the "Early" channels an "Late" channels respectively. The last vertical line indicate the end of the glitch.



interaction term *Period1* × *EarlyChannel* then captures the DiD effect of the glitch affected stocks (i.e., Nasdaq-listed stocks in the even-numbered channels) relative to unaffected stocks (i.e., both Nasdaq-listed stocks in the odd-numbered channels and NYSE-listed stocks). The results show that market liquidity of the affected stocks worsens significantly, with quoted spread increasing by 0.28 basis point or about 7.3% relative to pre-glitch level (0.28 / 3.84), trading volume falling by 25.88 shares or ≈ 39.6% (-25.88 / 62.18) relative to the pre-glitch level, depth on the NBBO decreasing by 8563 dollars (-85.86 / 425 ≈ 20.2% of pre-glitch level), dollar depth on the first 5 levels of NBBO drops by 367.08 thousand dollars (-367.08 / 2561.81 ≈ 26.5%). In contrast, coefficient on the interaction term *Period1* × *LateChannel* is insignificant for all liquidity measures, suggesting that the liquidity withdraw of the stocks in the early channels are indeed due to the glitch in their consolidated feeds.

Instead, the second DiD regression is run over the second period of the glitch, that is, when stocks of the late channels started to experienced a glitch as well. The coefficient on the interaction

Table 9. Difference-in-difference regression: SIP glitch on Jan 3, 2013.

(a) First glitch period. This table show regression results from the following difference-in-difference regression:

$$metric_{i,t} = \alpha_i + \beta Period1_{i,t} + \beta_1 Period1_{i,t} \times EarlyChannel_{i,t} + \beta_2 Period1_{i,t} \times LateChannel_{i,t} + \epsilon_{i,t}$$

where $metrc_{i,t}$ is the liquidity or trading metric of stock i in time interval t. α_i is the stock fixed effects. $Period1_{i,t}$ is dummy variable that equals one between the start of the data outage at early channels and late channels, that is, between 13:33:11 and 13:36:51, and equals zero otherwise. Standard errors are clustered at the stock level. $EarlyChannel_{i,t}$ is dummy variable that equals one if stock i belongs to one of the early channels, and equals zero otherwise. $LateChannel_{i,t}$ is dummy variable that equals one if stock i belongs to one of the late channels, and equals zero otherwise. Note for the first regression, I include stocks from all early, late and normal channels. The sample period is between 30 minutes before the outage started at early channels and the start of outage at late channels.

	RQS	RES	RRS	DollarVlm	DepthNBBO	Depth5lvl
Period1	−0.01	−0.03	0.08***	−2.21	30.94	−37.83
	(0.02)	(0.02)	(0.03)	(2.64)	(28.66)	(47.41)
Period1 x EarlyChannel	0.28***	−0.02	0.36***	−25.88***	−85.86**	−367.08***
	(0.11)	(0.09)	(0.10)	(4.70)	(38.57)	(129.29)
Period1 x LateChannel	−0.10	−0.05	0.06	−21.88	−35.72	−76.57
	(0.09)	(0.07)	(0.09)	(14.25)	(30.06)	(55.22)
R^2 (%)	0.16	0.01	0.08	0.04	0.14	1.04
N	55596	55596	55596	55596	55596	55596
Stock F.E.	Yes	Yes	Yes	Yes	Yes	Yes

(b) Second glitch period. This table show regression results from the following difference-in-difference regression:

$$metric_{i,t} = \alpha_i + \beta Period2_{i,t} + \beta_1 Period2_{i,t} \times LateChannel_{i,t} + \epsilon_{i,t}$$

where $metrc_{i,t}$ is the liquidity or trading metric of stock i in time interval t. α_i is the stock fixed effects. $Period2_{i,t}$ is dummy variable that equals one between the start of the data outage at late channels and the end of data outages, that is, between 13:36:51 and 13:51:15, and equals zero otherwise. Standard errors are clustered at the stock level. $LateChannel_{i,t}$ is dummy variable that equals one if stock i belongs to one of the late channels, and equals zero otherwise. Note that for the second regression, I only include stocks of the late channels and normal channels. The sample period is between the start of the outage at early channels and the end of all outages.

	RQS	RES	RRS	DollarVlm	DepthNBBO	Depth5lvl
Period2	−0.02	0.07**	0.05	1.40	−47.36***	−241.36***
	(0.03)	(0.03)	(0.04)	(3.11)	(16.37)	(37.81)
Period2 x LateChannel	0.43***	0.17	0.47***	−35.99*	−9.49	−20.95
	(0.13)	(0.13)	(0.16)	(18.53)	(23.32)	(122.22)
R^2 (%)	0.80	0.23	0.29	0.16	0.54	9.34
N	12265	12265	12265	12265	12265	12265
Stock F.E.	Yes	Yes	Yes	Yes	Yes	Yes

term *Period2* × *LateChannel* captures then the DiD effect of the stocks in the late channels compared to unaffected stocks (i.e., NYSE-listed stocks). The results show that quoted spread and effective spread both increase and trading volume falls for glitch affected stocks relative to unaffected stocks. However, depth on the NBBO and five levels within the NBBO do not decrease more for glitch affected stocks than unaffected stocks. It could again result from the illiquidity contagion from glitch affected stocks to unaffected stocks during the second period of the glitch event.

4 Conclusion

In this paper I study the role of the consolidated feed in the US equities market by examining exogenous events when it becomes faster due to a speed upgrade and when it is corrupted or unavailable due to a technical glitch. The unique structure of two consolidated feeds, one for Nasdaq-listed stocks and the other for NYSE-listed stocks, allows me to use difference-in-difference (DiD) analysis based on a matched sample of Nasdaq-listed stocks and NYSE-listed stocks for an identification. The results show that while a faster consolidated does not affect overall market liquidity, it has a significant impact on trader behavior. Specifically, after a speed upgrade to consolidated feeds for Nasdaq-listed stocks, low-latency activities increases in Nasdaq-listed stocks relative to NYSE-listed stocks, possibly due to more competition between fast traders and slow traders as the speed advantage of the former reduces . Last, when the consolidated feed becomes corrupted or unavailable due to technical glitches, market liquidity worsens significantly. The foregoing findings show that the consolidated feed matters and remains a crucial component of the current market data infrastructure. As consolidated feeds are the focus of the ongoing market structure reform agenda, more studies are needed to better understand the its role in today's market and assess the potential impacts of proposed new rules.

A Other summary statistics

Table A1. Summary statistics of the matched sample for the Nasdaq-SIP upgrade event on October 24. 2016: Exchange-specific metrics. This table shows the summary statistics of the exchange-specific liquidity and trading metrics by exchange.

Venue	Variable	N	Mean	SD	Min	10%	25%	50%	75%	90%	Max
Arca	RQS	7828	8.80	7.40	0.90	2.90	4.10	6.57	10.67	17.01	95.70
	RES	7828	1.27	0.94	0.22	0.55	0.72	1.00	1.46	2.33	12.87
	RRS	7828	-0.43	0.75	-7.10	-1.11	-0.68	-0.35	-0.09	0.18	22.59
	Vlm	7828	20.64	43.10	0.30	3.35	5.78	10.78	20.76	38.16	942.26
	DollarDepth	7828	33.36	53.68	2.85	7.56	10.81	16.50	28.94	70.68	500.15
	NBBOShr	7828	80.13	17.84	18.38	54.79	67.46	83.35	97.04	99.90	100.00
	ISOShr	7828	56.96	7.67	18.09	47.20	52.04	57.32	62.12	66.50	81.17
	Cancel/Trade	7752	30.54	15.38	4.64	15.07	20.18	27.37	37.27	49.26	298.78
	Order/Trade	7752	41.07	23.62	5.01	18.79	25.43	35.38	50.31	69.06	270.28
BZX	RQS	7828	10.93	9.33	0.92	2.97	4.35	7.93	14.06	23.72	67.23
	RES	7828	1.41	0.96	0.27	0.66	0.85	1.15	1.60	2.44	10.75
	RRS	7828	-0.54	0.69	-11.71	-1.21	-0.79	-0.47	-0.20	0.05	4.60
	Vlm	7828	15.76	29.25	0.38	2.68	4.44	8.28	16.65	31.51	573.20
	DollarDepth	7828	22.77	34.42	0.66	4.37	6.32	10.31	21.20	57.47	302.59
	NBBOShr	7828	68.67	26.03	5.17	30.65	47.68	70.19	95.90	99.88	100.00
	ISOShr	7828	55.76	7.36	24.21	46.25	51.10	56.13	60.84	65.01	81.33
	Cancel/Trade	7752	34.64	24.70	5.66	15.31	20.18	28.66	40.78	59.38	478.68
	Order/Trade	7752	45.16	31.43	6.62	20.31	26.55	36.96	52.84	77.50	470.04
EDGX	RQS	7828	23.60	38.11	0.90	2.87	4.46	10.13	25.93	55.02	417.47
	RES	7828	1.46	1.08	0.28	0.69	0.87	1.18	1.63	2.52	14.51
	RRS	7828	-0.61	0.82	-11.72	-1.40	-0.90	-0.51	-0.21	0.07	5.46
	Vlm	7828	19.48	44.16	0.19	2.36	4.15	8.76	18.89	37.91	740.55
	DollarDepth	7828	25.18	38.38	0.79	3.92	6.33	11.55	24.66	63.41	370.23
	NBBOShr	7828	69.28	26.94	6.26	30.94	45.74	72.02	97.65	99.89	100.00
	ISOShr	7828	54.58	7.64	23.16	44.56	49.51	54.78	59.85	64.07	78.90
	Cancel/Trade	7752	13.05	7.26	1.49	5.57	8.00	11.62	16.34	21.95	82.46
	Order/Trade	7752	17.44	10.18	2.83	7.87	10.81	15.23	21.20	28.94	143.72
Nasdaq	RQS	7828	5.58	3.83	0.87	2.34	3.13	4.45	6.82	10.19	35.34
	RES	7828	1.41	1.04	0.27	0.65	0.84	1.15	1.58	2.49	13.56
	RRS	7828	-0.65	0.71	-10.39	-1.31	-0.87	-0.54	-0.29	-0.07	2.49
	Vlm	7828	43.82	81.03	0.64	7.54	13.07	24.16	46.53	83.85	1854.37
	DollarDepth	7828	58.84	72.15	4.73	17.15	24.26	35.96	55.54	131.63	707.69
	NBBOShr	7828	93.78	8.89	32.55	81.39	91.14	97.85	99.67	99.97	100.00
	ISOShr	7828	55.67	6.90	24.29	46.70	51.18	55.82	60.37	64.36	80.97
	Cancel/Trade	7752	21.62	9.90	4.13	11.78	14.92	19.48	25.96	34.03	115.40
	Order/Trade	7752	32.52	18.77	5.83	16.42	21.04	27.91	38.32	53.15	446.34

Table A2. Summary statistics of the matched sample for the three pooled SIP glitch events. This table reports the summary statistics of the liquidity metrics for the pooled sample of SIP glitch events. All variables are defined above. The summary statistics are computed base on time-series averages over the sample period.

	N	Mean	SD	Min	10%	25%	50%	75%	90%	Max
RQS	355	4.39	3.18	0.76	2.06	2.62	3.65	5.17	7.31	38.99
RES	355	1.45	1.25	0.29	0.66	0.81	1.12	1.66	2.43	15.06
RRS	355	0.41	1.13	-2.01	-0.33	-0.13	0.17	0.55	1.34	12.31
DollarVlm	355	111.23	182.28	7.12	21.20	36.96	70.61	125.11	215.46	2770.43
DepthNBBO	355	227.10	681.63	32.39	52.31	66.07	94.34	157.56	360.68	11421.75
Depth5lvl	355	1366.47	3249.54	156.95	258.91	348.29	583.29	1302.83	2687.66	50766.24

Table A3. Summary statistics of the matched sample for the SIP glitch event on January 3, 2013: Period one. This table reports the summary statistics of the liquidity metrics for the pooled sample of the first period of the SIP glitch on January 3, 2013. All variables are defined above. The summary statistics are computed base on time-series averages over the sample period.

Variable	N	Mean	SD	Min	10%	25%	50%	75%	90%	Max
RQS	56016	3.84	3.43	0.41	1.56	2.05	2.87	4.38	7.05	39.14
RES	56016	1.48	1.60	-9.30	0.33	0.71	1.14	1.78	2.99	25.52
RRS	56016	0.74	2.25	-23.19	-1.45	-0.35	0.76	1.57	2.92	42.45
DollarVlm	56016	64.33	214.77	0.02	1.71	4.78	16.07	56.57	144.41	16916.27
DepthNBBO	56016	424.79	961.80	2.52	50.72	83.96	164.52	382.80	983.95	15503.97
Depth5lvl	56016	2543.85	4459.69	78.78	398.88	649.65	1340.49	2604.95	5348.70	56993.86

Table A4. Summary statistics of the matched sample for the SIP glitch event on January 3, 2013: Period two. This table reports the summary statistics of the liquidity metrics for the pooled sample of the second period of the SIP glitch on January 3, 2013. All variables are defined above. The summary statistics are computed base on time-series averages over the sample period.

Variable	N	Mean	SD	Min	10%	25%	50%	75%	90%	Max
RQS	24652	3.90	3.65	0.40	1.54	2.01	2.89	4.45	7.05	40.82
RES	24652	1.55	1.81	-15.57	0.34	0.73	1.16	1.87	3.08	29.59
RRS	24652	0.94	2.30	-20.62	-1.24	-0.04	0.86	1.73	3.02	29.30
DollarVlm	24652	53.62	155.35	0.01	1.56	4.49	13.83	48.66	125.85	8006.75
DepthNBBO	24652	385.61	919.49	0.74	48.94	81.96	157.37	356.17	784.49	13950.07
Depth5lvl	24652	2151.26	3985.41	58.56	361.35	547.53	1119.99	2270.07	4332.79	54208.06

B Liquidity and trading metrics around the Nasdaq-SIP speed upgrade.

Figure A1. Daily liquidity and trading metrics for NYSE-listed stocks versus Nasdaq-listed stocks around the Nasdaq-SIP speed upgrade on October 24, 2016. This figure plots the daily time-series of several liquidity and trading metrics for Nasdaq-listed stocks and a matched sample of NYSE-listed stocks. The vertical line represents the speed upgrade to the Nasdaq-SIP on October, 24, 2016. The sample period is from September 26 to December 1, 2016. *RQS*, *RES* stand for relative quoted spread and relative effective spread in basis point.*Depth* is dollar depth at NBBO in thousands. *Vlm* is dollar trading volume in millions. *Cancel/Trade* is the ratio of cancel order count to total trade count. *Order/Trade* is the ratio of order volume of add order messages to total trade volume. *PrcImpShr* is the trade volume that receives price improvement as a fraction of total dark trading volume. *ISOShare* is trade volume via inter-market sweep order (ISO) as a fraction of total trade volume.

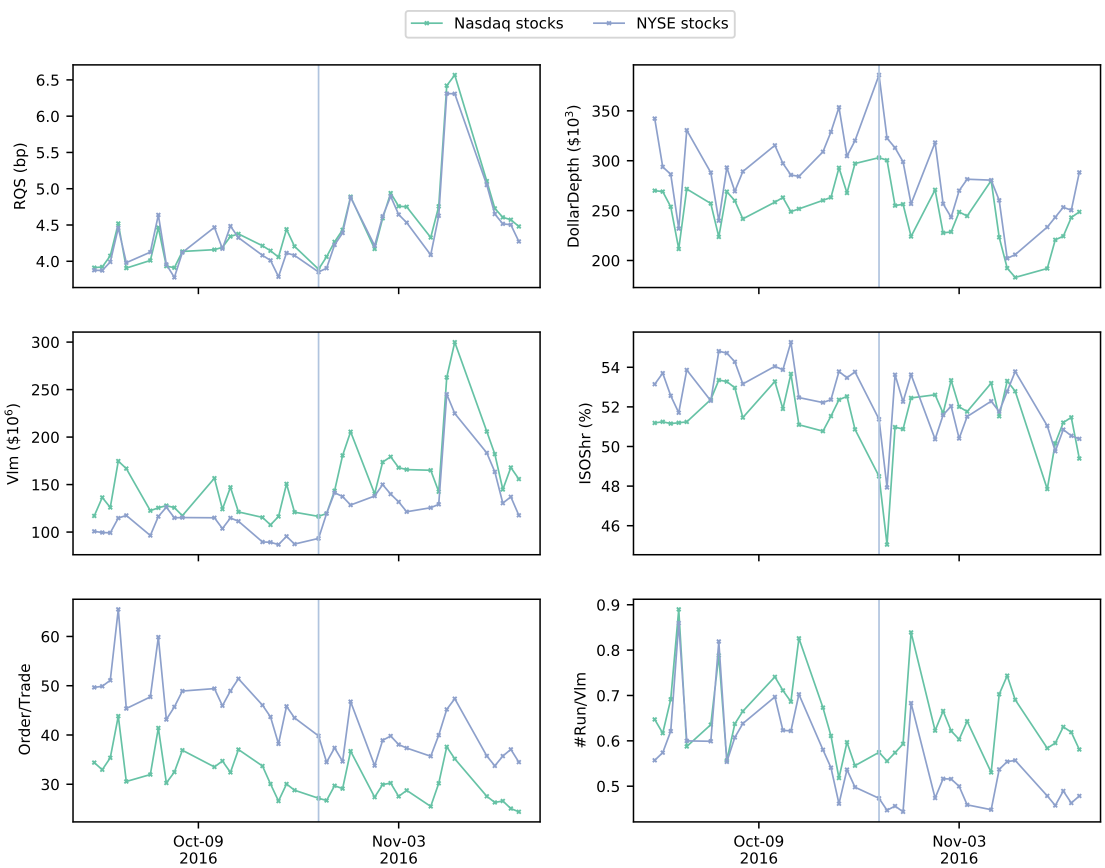
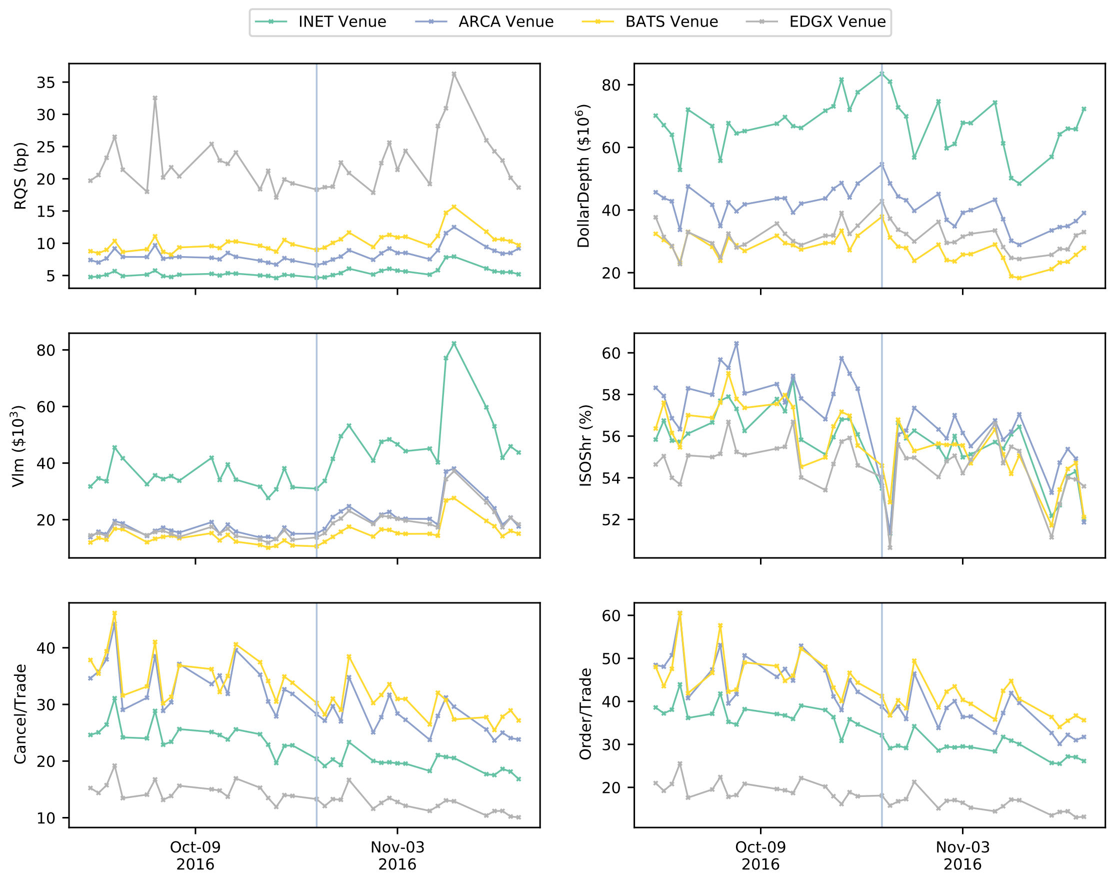


Figure A2. Daily liquidity and trading metrics on the Nasdaq exchange versus NYSE Arca and Bats around the Nasdaq-SIP speed upgrade on October 24, 2016. This figure plots the daily time-series of several liquidity and trading metrics on the Nasdaq exchange and NYSE Arca and Bats. The vertical line represents the speed upgrade to the Nasdaq-SIP on October, 24, 2016. The sample period is from September 26 to December 1, 2016. *RQS* stands for relative quoted spread in basis point. *Vlm* is dollar trading volume in millions. *Depth* is dollar depth at NBBO in thousands. *Cancel/Trade* is the ratio of cancel order count to total trade count. *Order/Trade* is the ratio of order volume of add order messages to total trade volume. *DepthShr* is the exchange's dollar depth at the NBBO as a fraction of total depth at the NBBO. *NBBOShr* is the exchange's time of being at the NBBO as a fraction of total trading hours.

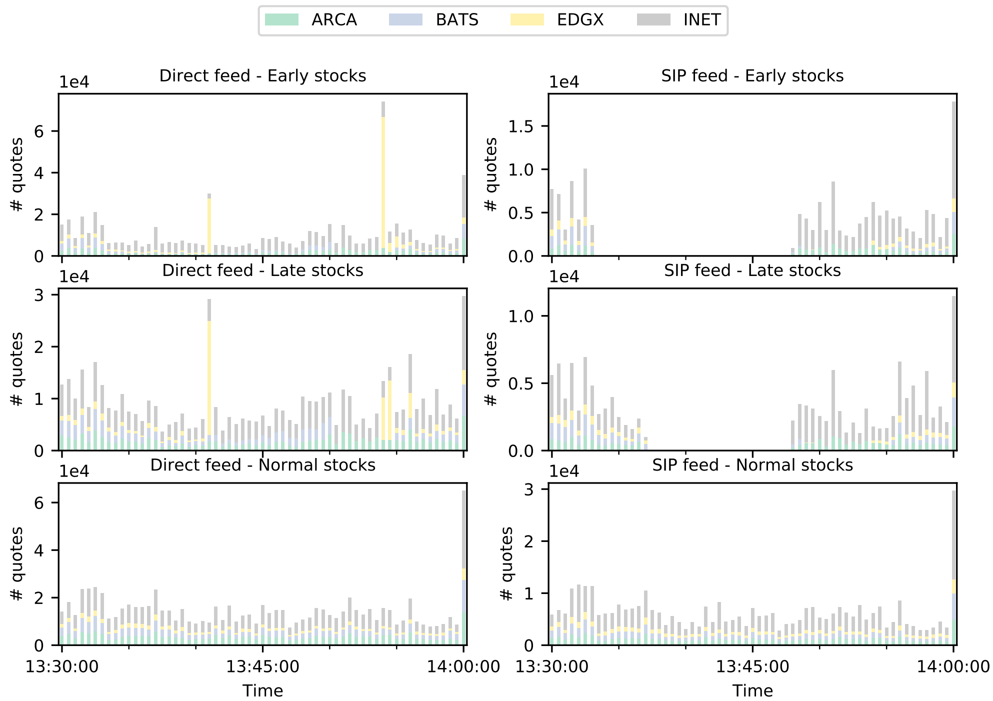


C SIP glitch events

- Figure A3 plots the number of trades and quote updates by exchange from direct feeds and consolidated feeds around the Nasdaq-SIP glitch event on January 3, 2013.
- Figure A4 plots the same metrics for the Nasdaq-SIP glitch event on October 30, 2014.
- Figure A5 plots the same metrics for the Nasdaq-SIP glitch event on August 22, 2019.

Figure A3. Nasdaq-SIP glitch event on January 3, 2013. This figure plots the number of trades and quote updates by exchange from direct feeds and consolidate feeds around the Nasdaq-SIP glitch event on January 3, 2013.

(a) Number of quotes.



(b) Number of trades.

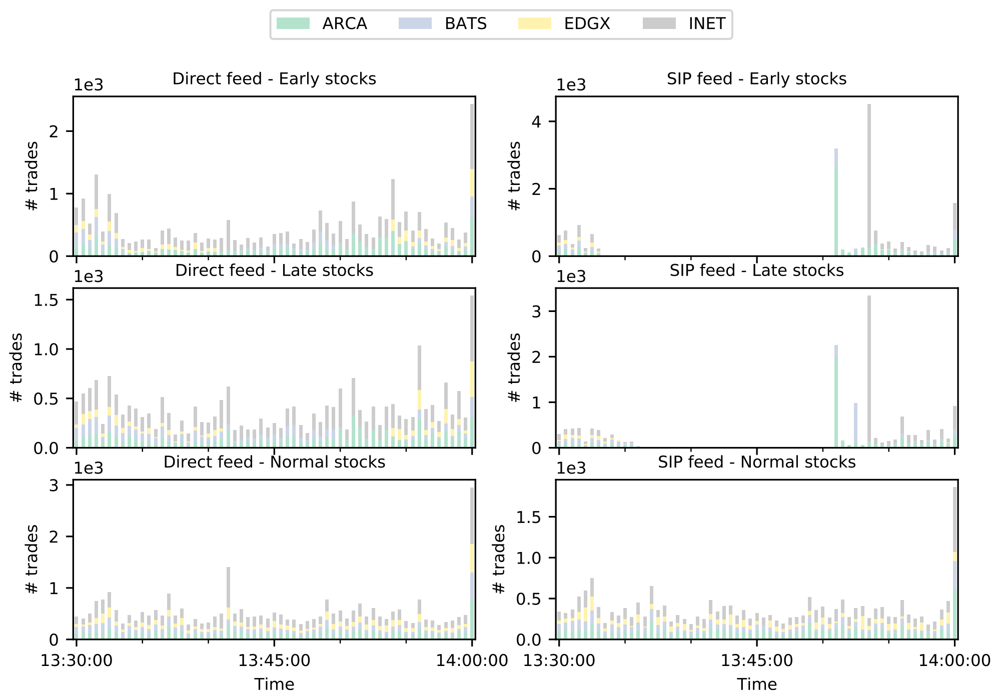
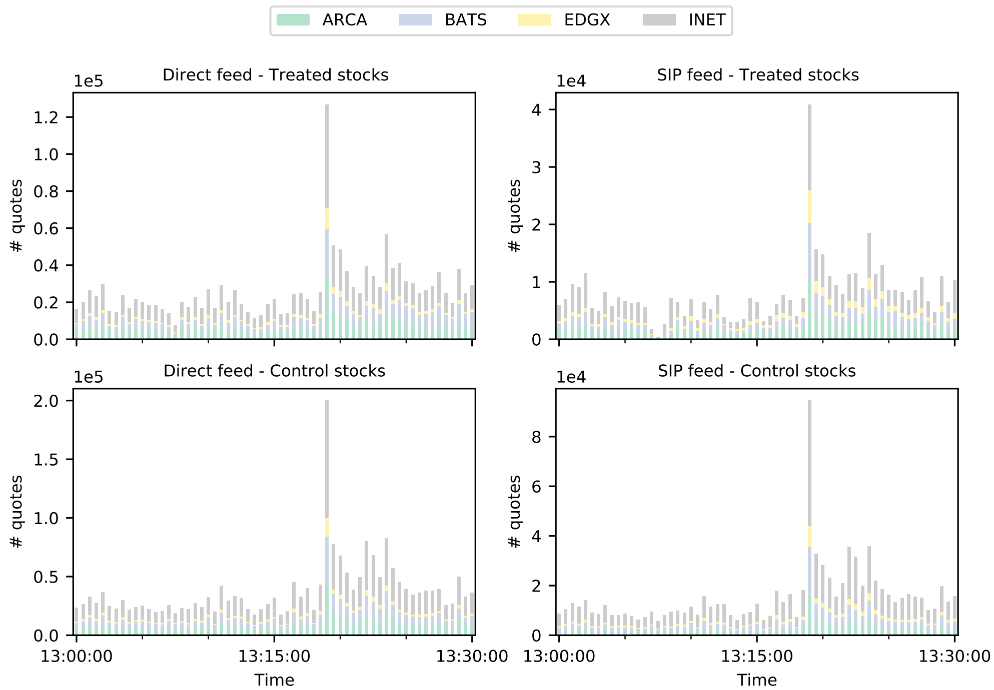


Figure A4. Nasdaq-SIP glitch event on October 30, 2014. This figure plots the number of trades and quote updates by exchange from direct feeds and consolidate feeds around the Nasdaq-SIP glitch event on October 30, 2014.

(a) Number of quotes.



(b) Number of trades.

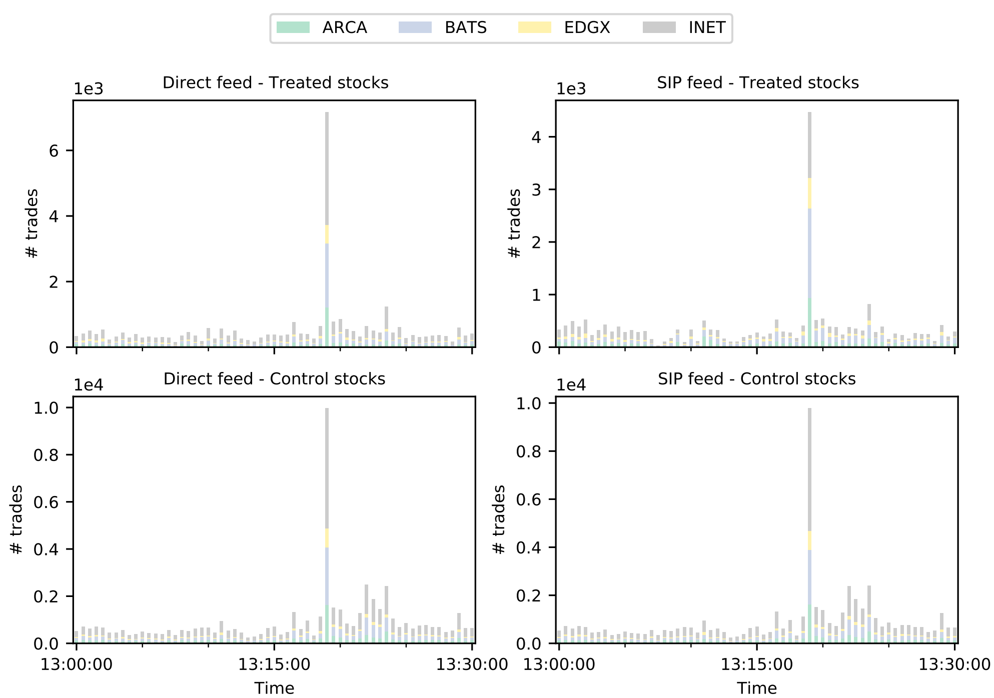
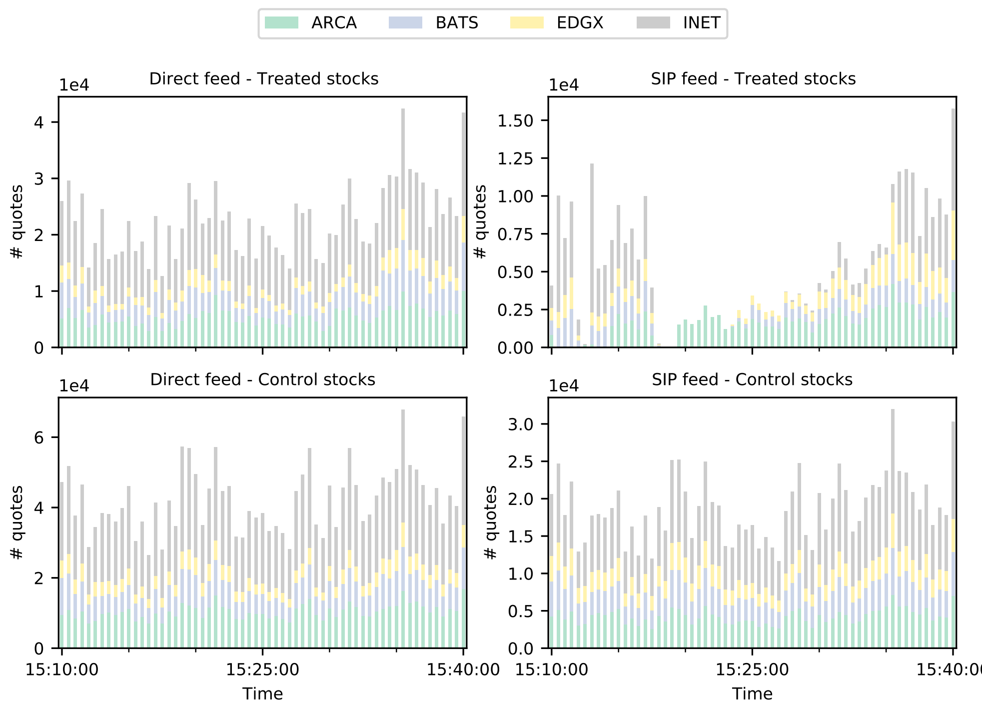
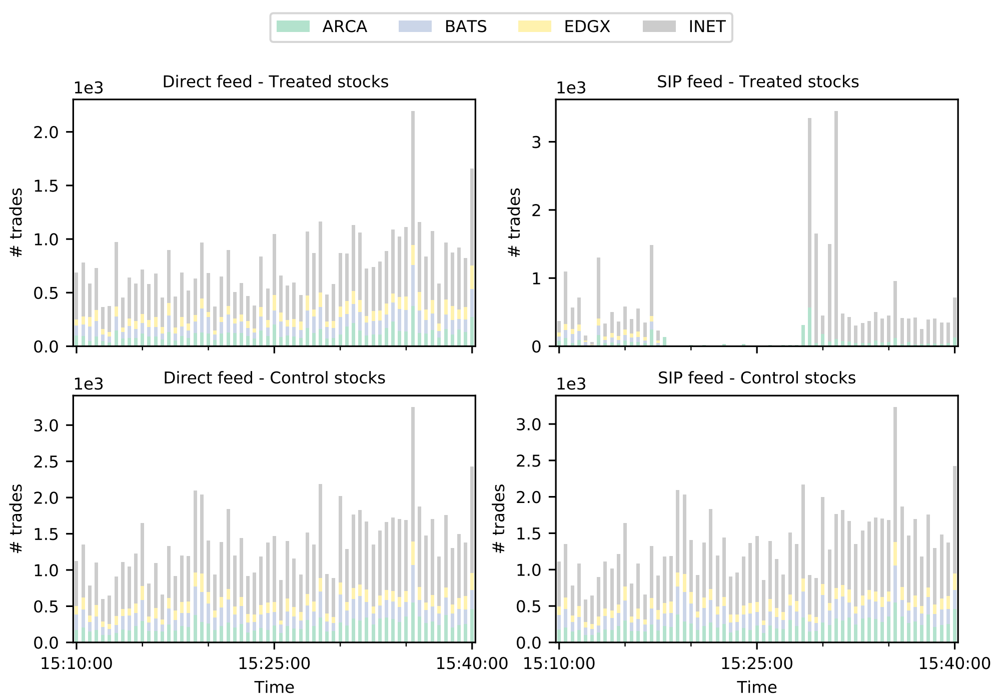


Figure A5. Nasdaq-SIP glitch event on August 12, 2019. This figure plots the number of trades and quote updates by exchange from direct feeds and consolidate feeds around the Nasdaq-SIP glitch event on August 12, 2019.

(a) Number of quotes.



(b) Number of trades.



References

Aldrich, Eric M. (2017). "The Flash Crash : A New Deconstruction". In: *SSRN Electronic Journal*, pp. 1–53.

Bartlett, Robert P. and Justin McCrary (2019). "How rigged are stock markets? Evidence from microsecond timestamps". In: *Journal of Financial Markets* 45.xxxx, pp. 37–60.

Battalio, Robert H., Shane A. Corwin, and Robert Jennings (2016). "Unrecognized odd lot liquidity supply". In: *Journal of Trading*, pp. 35–41.

Biais, Bruno, Thierry Foucault, and Sophie Moinas (2015). "Equilibrium fast trading". In: *Journal of Financial Economics* 116.2, pp. 292–313.

Boehmer, Ekkehart, Charles M. Jones, Xiaoyan Zhang, and Xinran Zhang (2021). "Tracking Retail Investor Activity". In: *The Journal of Finance*, jofi.13033.

Brogaard, Jonathan and James Brugler (2020). "Competition and Exchange Data Fees". In: *SSRN Electronic Journal* 1.801.

Brogaard, Jonathan, Björn Hagströmer, Lars Nordén, and Ryan Riordan (2015). "Trading Fast and Slow: Colocation and Liquidity". In: *Review of Financial Studies* 28.12, pp. 3407–3443.

Brogaard, Jonathan, Matthew C. Ringgenberg, and Dominik Rösch (2020). "Does Floor Trading Matter?" In: *SSRN Electronic Journal*.

Budish, Eric B., Peter Cramton, and John J. Shim (2015). "The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response". In: *The Quarterly Journal of Economics* 130.4, pp. 1547–1621.

Cespa, Giovanni and Thierry Foucault (2014). "Illiquidity contagion and liquidity crashes". In: *Review of Financial Studies* 27.6, pp. 1615–1660.

CFTC and SEC (2010). "Findings Regarding the Market Events of May 6, 2010". In: 20549.202, pp. 1–104.

Chakravarty, Sugato, Pankaj K. Jain, James Upson, and Robert Wood (2012). "Clean sweep: Informed trading through intermarket sweep orders". In: *Journal of Financial and Quantitative Analysis* 47.2, pp. 415–435.

Ding, Shengwei, John Hanna, and Terrence J. Hendershott (2014). "How slow is the NBBO? A comparison with direct exchange feeds". In: *Financial Review* 49.2, pp. 313–332.

Ernst, Thomas, Jonathan Sokobin, and Chester Spatt (2021). “The Value of Off-Exchange Data”. In: *Working Paper*.

Foucault, Thierry, Johan Hombert, and Ioanid Roşu (2016). “News Trading and Speed”. In: *Journal of Finance* 71.1, pp. 335–382.

Hagströmer, Björn and Lars Nordén (2013). “The diversity of high-frequency traders”. In: *Journal of Financial Markets* 16.4, pp. 741–770.

Hasbrouck, Joel (2019). “Price Discovery in High Resolution”. In: *Journal of Financial Econometrics*, pp. 1–36.

Hasbrouck, Joel and Gideon Saar (2013). “Low-latency trading”. In: *Journal of Financial Markets* 16.4, pp. 646–679.

Hendershott, Terrence J., Charles M. Jones, and Albert J. Menkveld (2011). “Does algorithmic trading improve liquidity?” In: *Journal of Finance* 66.1, pp. 1–33.

Hendershott, Terrence J., Marc Rysman, and Rainer Schwabe (2020). “Stock Exchanges as Platforms for Data and Trading”. In: *SSRN Electronic Journal*.

Hoffmann, Peter (2014). “A dynamic limit order market with fast and slow traders”. In: *Journal of Financial Economics* 113.1, pp. 156–169.

Jovanovic, Boyan and Albert J. Menkveld (2016). “Middlemen in Limit-Order Markets”. In: *SSRN Electronic Journal*.

Menkveld, Albert J. (2016). “The Economics of High-Frequency Trading: Taking Stock”. In: *Annual Review of Financial Economics* 8.1, pp. 1–24.

— (2018). “High-frequency trading as viewed through an electron microscope”. In: *Financial Analysts Journal* 74.2, pp. 24–31.

Menkveld, Albert J. and Marius Zoican (2016). “Need for Speed ? Exchange Latency and Liquidity”. In: *Review of Financial Studies* 30.4, pp. 1188–1228.

O’Hara, Maureen, Chen Yao, and Mao Ye (2014). “What’s not there: Odd lots and market data”. In: *Journal of Finance* 69.5, pp. 2199–2236.

Shkilko, Andriy and Konstantin Sokolov (2020). “Every Cloud Has a Silver Lining: Fast Trading, Microwave Connectivity, and Trading Costs”. In: *The Journal of Finance* 75.6, pp. 2899–2927.

Yao, Chen and Mao Ye (2018). “Why trading speed matters: A tale of queue rationing under price controls”. In: *Review of Financial Studies* 31.6, pp. 2157–2183.

Ye, Mao, Chen Yao, and Jiading Gai (2013). "The Externality of High Frequency Trading". In: *SSRN Electronic Journal*.